EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

24 March 2022
EMA/231935/2022
Committee for Medicinal Products for Human Use (CHMP)

Assessment report

Jakavi

International non-proprietary name: ruxolitinib

Procedure No. EMEA/H/C/002464/II/0053

Note

Variation assessment report as adopted by the CHMP with all information of a commercially confidential nature deleted.

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

Table of contents

List of abbreviations

ADR	Adverse Drug Reaction
AE	Adverse Event
AESI	Adverse Event of Special Interest
aGvHD	Acute Graft vs. Host Disease
alloSCT	Allogeneic Stem Cell Transplantation
ALL	Acute Lymphoblastic Leukemia
ALP	Alkaline Phosphatase
ALT	Alanine aminotransferase
AML	Acute Myeloid Leukemia
ANC	Absolute Neutrophils Count
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
ATG	Anti-Thymocyte Globulin
BID	bis in diem/twice a day
BALT	Bone Alkaline Phosphatase
BAT	Best Available Therapy
b.i.d.	bis in diem/twice a day
BOR	Best Overall Response
CAS	Cross over Analysis Set
cGvHD	chronic Graft vs. Host Disease
CIBMTR	Center for International Blood and Marrow Transplant Research
CMH	Cochrane–Mantel–Haenszel
CMV	Cytomegalovirus
CNI	Calcineurin Inhibitor
CV	Coefficient of variation
COVID-19	COVID-19
CR	Complete response
CRF	Case Report/Record Form; the term CRF can be applied to either EDC or Paper
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events
Ctrough	Minimum concentration
CYP3A	Cytochrome P450, family 3, subfamily A
CYP450	Cytochrome P450
EDTA	Ethylenediaminetetraacetic acid
DILI	Drug-Induced Liver Injury
DLI	Donor Lymphocyte Infusion
DMC	Data Monitoring Committee
DOR	Duration of Response
EBV	Epstein-Barr Virus
ECOG	Eastern Cooperative Oncology Group
ECP	Extracorporeal Photopheresis
eCRF	Electronic Case Report/Record Form
EDC	Electronic Data Capture
EFS	Event-Free Survival
EOS	End of Study
EOT	End of treatment
FACT-BMT	Functional Assessment of Cancer Therapy - Bone Marrow Transplantation

FAS	Full Analysis Set
FFS	Failure-Free Survival
GGT	Gamma-Glutamyl Transferase
GI	Gastro-Intestinal
GvHD	Graft vs. Host Disease
GvL	Graft vs. Leukemia
HLA	Human leukocyte antigen
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HCT	Hematopoietic Cell Transplantation
HDL	High Density Lipoprotein
HHV-6	Human Herpes Virus
HIV	Human Immunodeficiency Virus
HLA	Human Leukocyte Antigen
HR	Heart rate
HSCT	Hematopoietic stem cell transplantation
ICF	Informed consent form
ICH	International Council on Harmonization
ICU	Intensive Care Unit
IEC	Independent Ethics Committee
IFN	Interferon
IL	Interleukin
IRB	Institutional Review Board
IRT	Interactive Response Technology that includes Interactive Voice Response System and interactive Web Response System
ITT	Intent To Treat
JAK	Janus kinase
K-M	Kaplan-Meier
LDL	Low Density Lipoprotein
LC-MS/MS	Liquid Chromatography-tandem Mass Spectrometry
LLOQ	Lower limit of quantification
LMWH	Low molecular weight heparin
MCH	Mean Corpuscular Hemoglobin
MDS	Myelodysplastic Syndromes
MedDRA	Medical dictionary for regulatory activities
MMF	Mycophenolate mofetil
MF	Myelofibrosis
MPNs	Myeloproliferative Neoplasms
MR	Malignancy Relapse/Progression
MSC	Mesenchymal Stromal Cells
mTOR	Mammalian Target of Rapamycin
MTX	Methotrexate
NCA	Non-compartmental analysis
NIH	National Institutes of Health
NMSC	Non-Melanoma Skin Cancer
NRM	Non Relapse Mortality
NSAID	Nonsteroidal anti-inflammatory drug
o.d.	omnia die/once a day
OR	Odds ratio / Overall Response

ORR	Overall Response Rate
OS	Overall Survival
PAS	Pharmacokinetic analysis set
PGIC	Patient global impression of change
PGIS	Patient global impression of severity
PK	Pharmacokinetics
PPS	Per-Protocol Set
PV	polycythemia vera
PR	Partial response
PRO	Patient Reported Outcomes
PT	Preferred term
PTT	Partial thromboplastin time
PTY	Patient treatment years (instead of Subject treatment years)
QD	quaque die, once a day
QC	Quality control
QOL	Quality of Life
RAEB	Refractory anemia with excess blasts
RBC	Red blood cell
RDI	Relative Dose Intensity
ROW	Rest of the World
SAE	Serious Adverse Event
SAP	Statistical analysis plan
SCT	Stem cell transplantation
SD	standard deviation
SMQ	Standardized MedDRA Queries
SOC	System organ class
SOS	Sinusoidal Obstructive Syndrome
SR-cGvHD	Steroid refractory chronic Graft vs. Host Disease
TNC	Total nucleated dose
TNF	Tumor Necrosis Factor
TSS	Total symptom score
ULN	Upper limit of normal
US	United States of America
ULOQ	Upper limit of quantification
WBC	White Blood Cells

1. Background information on the procedure

1.1. Type II variation

Pursuant to Article 16 of Commission Regulation (EC) No 1234/2008, Novartis Europharm Limited submitted to the European Medicines Agency on 2 February 2021 an application for a variation.

The following changes were proposed:

Variation requested		**Type**	**Annexes affected**
C.I.6.a	C.I.6.a - Change(s) to therapeutic indication(s) - Addition of a new therapeutic indication or modification of an approved one	Type II	I and IIIB

Extension of indication to include treatment of patients with GvHD aged 12 years and older who have inadequate response to corticosteroids or other systemic therapies for Jakavi; as a consequence, sections 4.1, 4.2, 4.4, 4.5, 4.8. 5.1 and 5.2 of the SmPC are updated. The Package Leaflet is updated in accordance. Version 13.0 of the RMP has also been submitted. In addition, the Marketing authorisation holder (MAH) took the opportunity to update the list of local representatives for The Netherlands in the Package Leaflet.

The requested variation proposed amendments to the Summary of Product Characteristics and Package Leaflet and to the Risk Management Plan (RMP).

Information on paediatric requirements

Pursuant to Article 8 of Regulation (EC) No 1901/2006, the application included EMA Decisions P/0172/2021 (EMEA-000901-PIP03-16-M02) and P/0384/2019 (EMEA-000901-PIP04-17-M01) on the agreement of a paediatric investigation plan (PIP).

At the time of submission of the application, the P/0172/2021 (EMEA-000901-PIP03-16-M02) and P/0384/2019 (EMEA-000901-PIP04-17-M01) were not yet completed as some measures were deferred.

Information relating to orphan market exclusivity

Similarity

Pursuant to Article 8 of Regulation (EC) No. 141/2000 and Article 3 of Commission Regulation (EC) No 847/2000, the MAH did not submit a critical report addressing the possible similarity with authorised orphan medicinal products because there is no authorised orphan medicinal product for a condition related to the proposed indication.

Scientific advice

Scientific Advice was sought at the CHMP on clinical development for the Phase III study D2301 protocol in chronic GvHD (Procedure No: EMEA/H/SA/1155/2/2016/II).

1.2. Steps taken for the assessment of the product

The Rapporteur and Co-Rapporteur appointed by the CHMP were:

Rapporteur: Filip Josephson Co-Rapporteur: Paula Boudewina van Hennik

Timetable	Actual dates
Submission date	2 February 2021
Start of procedure:	26 April 2021
CHMP Co-Rapporteur's preliminary assessment report circulated on	25 May 2021
CHMP/PRAC Rapporteurs preliminary joint assessment report circulated on	26 May 2021
Updated PRAC Rapporteur's assessment report circulated on	3 June 2021
PRAC RMP advice and assessment overview adopted by PRAC on	10 June 2021
Updated CHMP Rapporteur's assessment report circulated on	18 June 2021
Request for supplementary information adopted by the CHMP on	24 June 2021
Summary report of the inspection as issued on	22 December 2021
MAH's responses submitted to the CHMP on	14 October 2021
CHMP/PRAC Rapporteurs preliminary 'joint assessment report on the MAH's responses circulated on	16 November 2021
Updated PRAC Rapporteur's assessment report on the MAH's responses circulated on	26 November 2021
PRAC RMP advice and assessment overview adopted by PRAC on	2 December 2021
Updated CHMP Rapporteur's assessment report on the MAH's responses circulated on	9 December 2021
Request for supplementary information adopted by the CHMP on	16 December 2021
MAH's responses submitted to the CHMP on	20 January 2022
CHMP/PRAC Rapporteurs preliminary joint assessment report on the MAH's responses circulated on	23 February 2022
PRAC RMP advice and assessment overview adopted by PRAC on	10 March 2022
Updated CHMP Rapporteur's assessment report on the MAH's responses circulated on	17 March 2022
CHMP opinion	24 March 2022

2. Scientific discussion

2.1. Introduction

Ruxolitinib (Jakavi®/Jakafi®, INC424, INCB018424 phosphate) is an oral selective inhibitor of the Janus kinases (JAKs) JAK1 and JAK2. Ruxolitinib (Jakavi®) is currently indicated in EU and more than 100 countries for the treatment of disease-related splenomegaly or symptoms in adult patients with PMF, PPV-MF or PET-MF and for the treatment of adult patients with polycythemia vera who are resistant to or intolerant of hydroxyurea.

The target indication subject to this Type II variation, is to extend the indications for ruxolitinib, as follows:

- Jakavi is indicated for the treatment of patients with GvHD aged 12 years and older who have inadequate response to corticosteroids or other systemic therapies.

2.1.1. Problem statement

Allogeneic hematopoietic stem cell transplantation (alloSCT) is a well-established procedure for the treatment of malignant and non-malignant hematological diseases (Copelan 2006). However, despite the curative potential of alloSCT, graft versus host disease (GvHD) is a major barrier to the efficacious outcome.

GvHD is an immunologically mediated, multi-organ disorder that occurs when donor-derived immune cells recognize the transplant recipient cells, organs and tissues as non-self, thereby initiating an adverse immune reaction leading to tissue damage, organ failure, or even death (Jagasia et al 2018, Greinix et al 2011, Ferrara et al 2009). GvHD It is the major cause of transplant-related morbidity and mortality, affecting up to 70% of HSCT recipients and accounting for 21-31% of deaths post-alloSCT among patients who received human leukocyte antigen (HLA)-matched sibling and 31-40% of deaths in patients receiving unrelated donor transplants (D'Souza and Fretham 2019, Jagasia et al 2018, Hill et al 2018, Zeiser and Blazar 2017, Pavletic and Fowler 2012).

GvHD is categorized into two main clinical forms namely acute GvHD (aGvHD) and chronic GvHD (cGvHD); although patients may also have disease characteristics of both, acute and chronic GvHD. Overlap syndrome, although infrequent, may occur with the presence of one or more features of aGvHD in patients with diagnosis of cGvHD or may develop clinical features of aGvHD after diagnosis of cGvHD (Jagasia et al 2015).

Acute GvHD

Acute GvHD is mainly characterized by mature donor T cell-mediated inflammatory disease (Hill et al 2018, Zhang et al 2006). It usually presents early after engraftment, with a median time to onset (grade II-IV) of 20-25 days (Axt et al 2019, Flowers et al 2011). The clinical manifestations are seen primarily in three organs: the skin (maculopapular erythematous skin rash, erythroderma), the liver (cholestasis, hyperbilirubinemia, and/or jaundice), and the lower and upper gastrointestinal tract (nausea, abdominal pain, vomiting, anorexia with weight loss, secretory diarrhea, GI bleeding and/or ileus) (Schoemans et al 2018, Harris et al 2016, Dignan et al 2012, Jacobsohn and Vogelsang 2007).

The overall severity of aGvHD is graded from I (mild) to IV (life-threatening) according to the degree of involvement of the individual organs (Schoemans et al 2018). The diagnosis and clinical staging are assessed using a comprehensive and systematic approach for the determination of GvHD onset, confidence in the attribution of symptoms to aGvHD, and quantification of clinical severity of aGvHD in each target organ (The MAGIC criteria, Harris et al 2016). The extent of individual organ staging and overall grade of aGvHD is an important prognostic factor.

Among all patients who had alloSCT, the occurrence of aGvHD ranges from 30% to 50%, with 14% to 36% developing severe (grade III to IV) aGvHD (Malard et al 2020, Zeiser and Blazer 2017). The incidence of grade II-IV aGvHD in pediatric patients (2-12 years old) is lower than in adults, but adolescent patients (13-17 years old) can present higher rates, closer to that seen in adults (Gatza et al 2020, Qayed et al 2018).

The occurrence and severity of aGvHD depends on various factors including donor type (i.e., matched or unmatched, related or unrelated), older patient/donor age, gender disparity, multiparous female

donors, intensity of transplant conditioning regimen, and absence of or suboptimal GvHD prophylaxis (Nassereddine et al 2017, Ringden et al 2009).

The likelihood of response to treatment decreases with increasing disease severity (Hill et al 2018). The overall aGvHD grade has major impact on the survival post HSCT. Adult and pediatric patients with grade III-IV aGvHD have a high mortality risk with a 2-year survival rate of 27-35% (Khoury et al 2017).

Acute GvHD is one of the most consistently reported risk factors for development of cGvHD. It is estimated that approximately 30-50% of the patients with aGvHD will develop cGvHD despite the treatment received (Pagliuca et al 2021, Ringden et al 2018).

Chronic GvHD

According to the current knowledge of cGvHD pathophysiology it begins with activation of host antigen-presenting cells (APC) expressed by damaged tissues and/or pathogens (Dhir et al 2014). Activated host APC then present host antigens to donor immune cells, leading to donor T-cell proliferation and inflammatory cytokine production. Cytokine dysregulation has also been implicated through observations that high levels of interleukin (IL)-1β, IFNγ, and tumor necrosis factor (TNF)-α are associated with more severe cGvHD (Socie and Ritz 2014). These inflammatory cytokines then recruit and induce proliferation of additional immune effector cells, thereby perpetuating an adverse cycle of alloreactive tissue injury and inflammation (Paczesny et al 2010).

Chronic GvHD is a major long-term complication after alloSCT (Jagasia et al 2015), occurring most frequently after 100 days post-transplant with a median time to onset reported as 162 days post-transplant (Flowers et al 2011). While aGvHD is mainly a mature donor T cell-mediated inflammatory disease, cGvHD is characterized by the activation of complex signaling pathways in both T and B cells, reduced levels of circulating regulatory B cells (Bregs) and CD4+ Tregs (Hill et al 2018, Zhang et al 2006).

Chronic GvHD usually involves not only the epithelial target tissues affected in classic aGvHD (GI tract, liver, skin,) but also additional organ systems including lungs, muscles, fascia, joints, genitalia, eyes, and nails, (Jagasia et al 2015, Dhir et al 2014, Greinix et al 2011). The signs and symptoms include rash, raised or discoloured skin, thickening or tightening of skin, dry mouth, yellow discoloration of skin/eyes dry eyes, shortness of breath, weight loss, difficulty swallowing, fatigue, and muscle weakness (Pavletic et al 2006).

Chronic GvHD is classified into mild, moderate, and severe based on degree and number of organs or sites involved according to standard criteria (Schoemans et al 2018, Jagasia et al 2015).

Among patients who undergo alloSCT, cGvHD occurs in 30% to 70% of patients. The occurrence of cGvHD varies depending on the donor type, with 40% of recipients from matched sibling donor (Zeiser and Blazar 2017a, Lazaryan et al 2016). Approximately 30% of cGvHD are de novo without any preceding aGvHD (Lee 2017). The rates of cGvHD are lower in paediatric and adolescent patients as compared to adults (Dhir et al 2014, Flowers et al 2011, Kernan et al 1993).

Chronic GvHD is a leading cause of non-relapse mortality and morbidity in patients surviving more than 2 years after transplantation. Chronic GvHD adversely affects physical and functional well-being as well as quality of life of most of the patients who are otherwise cured for their underlying disease after HSCT (Lee 2017, Arai et al 2015). The risk factors including grafting with growth factor mobilized blood cells and use of female donor for male recipients and unrelated donors associated with chronic GvHD were not changed after adjustment for prior aGvHD, suggesting that chronic GvHD is not simply an evolution of preceding acute GvHD (Flowers et al 2011).

It is important to realize that the NIH diagnostic criteria (Jagasia 2015) were devised for clinical trials to ensure that study participants had unequivocal chronic GVHD. Many patients with signs and symptoms encountered in practice will not meet the NIH diagnostic criteria for chronic GVHD but nevertheless have active allo-immunity requiring systemic immunosuppression to improve symptoms and prevent ongoing organ damage.

2.1.2. Management

There are no optimal or standard preventive methods for GvHD defined and treatments vary across institutions. Prophylaxis is mainly based on suppressing the donor T-cell function using immunosuppression regimens (calcineurin inhibitors (CNI), methotrexate (MTX), mycophenolate mofetil (MMF), anti-thymocyte globulin (ATG), rituximab or T-cell depletion) (Hamilton 2018, Ruutu et al 2014), but often at the risk of underlying disease relapse or rejection (Ruutu et al 2012).

Treatment for GvHD is based on the severity of the disease and the number of organs affected (please refer to aGvHD target organ staging below (Please refer to assessor's comment, section 4.5.3.1.1. for aGcHD and the assessor's comment, section 4.5.5.1.4 for cGvHD). Topical therapies including corticosteroids or CNI are recommended for patients with grade I aGvHD and mild cGvHD, whose disease is localized to the skin (Nassereddine et al 2017, Garnett et al 2013). Systemic treatments with corticosteroids are the standard first-line treatment for grade II to IV aGvHD and moderate to severe cGvHD.

Management Acute GvHD

In aGvHD, high dose systemic corticosteroids (methylprednisolone 2 mg/kg/day or prednisone 2.0-2.5 mg/kg/day) is the standard initial treatment of grade II to IV aGvHD (Penack et al 2020, Ruutu et al 2014, Martin et al 2012).

However, approximately 50% of patients with grade II to IV aGvHD do not show adequate response to corticosteroids and often become steroid resistant/refractory or fail to taper corticosteroids (Schoemans et al 2018, Jamil and Mineishi 2015). In addition, less than 50% of responding patients treated with steroids present a sustained response (Garnett et al 2013). Patients with aGvHD refractory to steroid therapy are at a high mortality risk with an estimated 2-year survival rate below 20% (Malard et al 2020).

Management chronic GvHD

In cGvHD, systemic steroids (prednisone 1 mg/kg) with or without addition of CNIs are the recommended first-line therapy for patients with moderate to severe cGvHD (Penack et al 2020). Approximately 50% to 60% of patients do not respond or have inadequate control of disease with steroid treatment and require addition of another systemic therapy or fail to taper corticosteroids (Axt et al 2019, Inamoto et al 2014, Garnett et al 2013). Among patients who respond to treatment, responses were durable in 20% to 40% patients, and remaining patients were considered either resistant or refractory to steroids (Mawardi et al 2019, Garnett et al 2013). The estimated median duration of treatment with immunosuppressive therapy was 2 to 3 years (Jamil and Mineishi 2015). The long-term immunosuppressive treatment along with the disease associated immunodeficiency further increases the risk of serious and life-threatening infections.

For patients with cGvHD who do not respond to steroids or are unable to taper steroids, the prognosis remains poor with a 5-year survival rate of 50 to 70%, necessitating the addition of other agents (Mawardi et al 2019, Wolff D et al 2011).

Tapering of corticosteroids is recommended in patients with clinical improvements to minimize the risk of infection and other toxicities. First-line treatment with a combination of steroids and other agents had failed so far to improve the outcomes and are associated with increased mortality or only sub-optimal response rates. Adding additional therapy to prophylaxis and/or first line therapy is generally recommended for patients who fail to respond to corticosteroids, do not tolerate or fail to taper corticosteroids. The term "steroid intolerance" has not been formally validated but refers to the emergence of unacceptable toxicity (e.g. uncontrolled infections, avascular necrosis, arterial hypertension, diabetes mellitus, myopathy, osteoporosis, etc.) attributed to corticosteroids, as evaluated by a healthcare professional.

Due to the lack of large scale, positive randomized prospective studies to compare the efficacy and safety of second-line therapy for GvHD, no standard second-line treatment is defined in the EU. However, the most common used second-line agents recommended in both acute and chronic GvHD include ruxolitinib, extracorporeal photopheresis (ECP), low-dose MTX, MMF, mTOR inhibitors (everolimus or sirolimus), or infliximab. These agents may be used alone or in combinations with steroids (Penack et al 2020, Ruutu et al 2014, Wolff et al 2011). Anti-thymocyte globulin (ATG), etanercept, and mesenchymal stromal cells (MSC) are also recommended for use as second line treatment of aGVHD. Additional treatments in cGvHD include rituximab, imatinib, pentostatin, and ibrutinib. A need exists for novel therapies in this hard-to-treat population.

Ruxolitinib was first approved (treatment of MF) under the trade name of Jakafi® in Nov-2011 in the USA. Ruxolitinib is since 2019 registered for the treatment of SR-aGvHD in the USA, based on the data from single-arm, multi-centre, Phase II Study 18424-271 (REACH 1; supportive study for this submission; hereafter referred to as Study 271) (Jagasia et al 2020). Furthermore, the BTK inhibitor ibrutinib, is also approved in the USA and Canada since 2017, for the treatment of patients with cGvHD who had failed one to three lines of prior systemic therapy, based on the results from a Phase Ib/II single arm study (n=42).

2.1.3. About the product

Ruxolitinib (Jakavi®/Jakafi®, INC424, INCB018424 phosphate) is a potent and selective inhibitor of Janus Kinases (JAKs), JAK1 and JAK2, mutated JAK2V617F, with moderate and minimal inhibitory activity against TYK2 and JAK3. Ruxolitinib interferes with the signalling of a number of cytokines and growth factors that are important for hematopoiesis and immune function. The JAK-STAT pathway plays a critical role in cytokine signalling and the development of several immune cell types, leading to T-cell proliferation, tissue damage and is a promising target for GvHD treatment (Schwartz et al 2016).

Inhibition of JAK1/2 signalling results in reduced proliferation of donor effector T cells, suppression of pro-inflammatory cytokine production in response to alloantigen, as well as impairment of antigen presenting cells, based on in vitro and in mouse models. Data from murine and mice models suggest that graft-versus-leukaemia effect of alloreactive T cells was maintained after ruxolitinib administration.

2.1.4. The development programme/compliance with CHMP guidance/scientific advice

Scientific Advice was sought in the EU for the Phase III study D2301 protocol in chronic GvHD (Procedure No.: EMEA/H/SA/1155/2/2016/II). The CHMP acknowledged that a placebo-controlled study would not be possible considering the differences between the best available therapies (BAT) allowed in the comparator arm. The listing of acceptable BAT was endorsed but it was noted that the many options will add to the heterogeneity of the control group. The CHMP suggested to consider

limitation of cross-over to 'early-progressors'. Regarding the proposed study population, it was noted that adolescents aged ≥12 years might significantly differ in terms of underlying disease, prior anti-neoplastic treatment, type of donor and stem cell source used and overall immunotolerance with unknown impact on study outcomes. Moreover, preclinical safety data in juvenile rats suggested bone effects of ruxolitinib with unknown consequences for safety in children. Next, it was advised to restrict the trial participation to true steroid-refractory patients according to NIH consensus on criteria for clinical trials in cGvHD, i.e. excluding steroid-dependent GvHD patients that are not able to taper the corticosteroid dose. Alternatively, steroid-refractoriness/steroid-dependency according to NIH definition should at least be considered for inclusion as stratification factor. With regard to the proposed endpoints, CHMP stated that clinically significant and statistically compelling results in favour of ruxolitinib in terms of ORR and failure free survival (FFS), if supported by consistent findings in relevant secondary endpoints, without detrimental effects on OS and the cumulative incidence of relapse of the underlying malignancy, could be sufficient to conclude on clinical benefit. The use of the Lee Chronic GVHD Symptom Scale instrument (Lee 2002), FACT-BMT, and EQ-5D to characterize patients reported outcomes and symptom burden improvement was accepted.

Pre-submission meetings with Rapporteur and Co-Rapporteur were held in November 2020.

A Paediatric Investigation Plan (PIP) has been agreed with the Paediatric Committee for aGvHD (EMEA-000901-PIP03-16-M01) and cGvHD (EMEA-000901-PIP04-17-M01).

2.1.5. General comments on compliance with GCP

The clinical studies included in this application were, as claimed by the applicant, conducted in full compliance with current Good Clinical Practices (ICH E6). Nevertheless, during assessment a GCP inspection has been performed due to the concerns raised by the high numbers of protocol deviations. Further details and impact of the inspection is provided in section 2.5.5 discussion on clinical efficacy. Overall, according to the inspectors, the observed findings are unlikely to have a significant impact on data integrity within the inspected clinical trials. The inspection team did not identify any restrictions on the usability of the reported trial data, and therefore it was the recommendation of the inspectors that the data of the REACH-2 and REACH-3 clinical trials could be used for evaluation and assessment of the application. The Committee/Rapporteurs shares this view and therefore concludes that the findings are unlikely to have had any significant impact on the benefit-risk balance with regard to both aGvHD and cGvHD.

2.2. Quality aspects

This variation does not include any specific quality variation application and the already approved formulations, 5, 10, 15 and 20 mg tablets for adults, are proposed for the paediatric population from 12 years of age. The 5 and 10 mg tablets are round curved tablets, 7.5 and 9.3 mm in diameter respectively. The 15 and 20 mg tablets are ovaloid or elongated tablets of approximately 15 x 7 mm and 17 x 7 mm respectively. The tablets contain the excipients microcrystalline cellulose, magnesium stearate, colloidal anhydrous silica, sodium starch glycolate, povidone, hydroxypropylcellulose and lactose monohydrate.

Suitability of the formulation for the paediatric population from 12 years of age.

Since the variation concerns an introduction of a paediatric population, the suitability of the proposed formulation in the proposed age group should be addressed, in line with the Guideline on pharmaceutical development of medicines for paediatric use EMA/CHMP/QWP/805880/2012 Rev.2. No such justification has been provided but this can be accepted since the proposed adult formulation only

contains commonly used excipients in amounts for which no safety issues are foreseen in the proposed target age group. The proposed adult formulation is considered acceptable from a quality point of view.

2.3. Non-clinical aspects

With this submission, the MAH has provided data pharmacology non clinic studies with ruxolitinib in acute and chronic Graft versus Host Disease (GvHD). All other aspects in of the nonclinical program remain unchanged, since the submission of the original dossier.

2.3.1. Pharmacology

Ruxolitinib oral administration in an MHC-mismatched allo-HSCT mouse model resulted in the decreased expression of inflammatory cytokines, reduced immune-cell infiltration in diseased colon tissue, reduced GvHD scores and body weight loss, preserved engraftment, and improved survival. Ruxolitinib inhibited pSTAT3 and pSTAT5 phosphorylation in T cells and in the inflamed colon tissue, consistent with JAK-STAT pathway inhibition. Ruxolitinib was efficacious in ameliorating disease severity in steroid-refractory mice with acute GvHD, leading to significant improvements in percent body weight loss, GvHD score, and increased survival.

Additionally, ruxolitinib was efficacious in ameliorating disease severity in a MHC-matched, miHA-mismatched chronic GvHD mouse model, leading to reduced GvHD scores, weight gain, significant survival benefits, improved skin integrity, and reduced incidence of skin and lung inflammation.

2.3.2. Ecotoxicity/Environmental risk assessment

For this new indication, the MAH has calculated a PEC_{sw} incorporating the previously approved indications.

Total $PEC_{surfacewater}$

Indication	$PEC_{SURFACEWATER}$ ($\mu g.L^{-1}$)
Graft versus Host Disease	0.00047
Polycythemia vera	0.0075
Myelofibrosis	0.000625
Total $PEC_{SURFACEWATER}$	**0.008**

It is concluded that PECsw remains below the trigger value and no further steps are taken. Previous assessment has demonstrated ruxolitinib not be a PBT substance.

2.3.3. Conclusion on the non-clinical aspects

The presented non clinical data have provided proof of concept support for the clinical development of ruxolitinib in treatment of acute and chronic GvHD. The clinical data further developed has then provided the necessary safety and efficacy data to support the extension of indication. No further non-clinical data have been provided which is considered acceptable. The updated data submitted in this application do not lead to a significant increase in environmental exposure further to the use of ruxolitinib in the new indication.

2.4. Clinical aspects

2.4.1. Introduction

Two pivotal, open label Phase III studies, REACH2 [CINC424C2301] for aGvHD and REACH3 [CINC424D2301] for chronic graft-versus-host disease (cGvHD) are conducted to support the submission. In both studies, a dose of ruxolitinib 10 mg orally b.i.d was used. The 10 mg b.i.d dose was the same in adolescents as in adult patients.

GCP

The Clinical trials were performed in accordance with GCP as claimed by the MAH.

The MAH has provided a statement to the effect that clinical trials conducted outside the community were carried out in accordance with the ethical standards of Directive 2001/20/EC.

- Tabular overview of clinical studies included in the submission and their status

Study/ Design	Number of patients enrolled/treated	Study population/ Primary endpoint	Status/ Cut-off date
Acute GvHD			
Study C2301 (Registration study)/ Phase III, open-label, randomized, multicenter study comparing the efficacy and safety of ruxolitinib vs. BAT (as selected by the Investigator) added to the patient's immunosuppressive regimen	N=309[1]/302 Ruxolitinib: n=154/152 BAT: n=155/150 [1] Ruxolitinib after cross-over: n= 49	Adults and adolescents (≥ 12 years old) with grade II to IV SR-aGvHD after alloSCT/ ORR at Day 28	Ongoing/ Primary analysis cut-off date: 25-Jul-2019 Secondary analysis cut-off date:06-Jan-2020
Study 271 (Supportive study) Phase II, open-label, single-cohort, multicenter study evaluating the efficacy and safety of ruxolitinib in combination with corticosteroids	N=71	Adult and adolescents (≥ 12 years old) with grade II to IV SR-aGvHD after alloSCT/ ORR at Day 28	Completed/ Final analysis cut-off date: 05-Jun-2019
Chronic GvHD			
Study D2301 (Registration study) Phase III, open-label, randomized, multicenter study comparing the efficacy and safety of ruxolitinib vs. BAT (as selected by the Investigator) added to the patient's immunosuppressive regimen	N=329/323 Ruxolitinib: n=165/165 BAT: n=164/158 Ruxolitinib after cross-over: n= 61	Adults and adolescents (≥ 12 years old) with moderate or severe SR-cGvHD after alloSCT/ ORR at Cycle 7 Day 1.	Ongoing/ Interim analysis cut-off date: 09-Jul-2019 Primary analysis cut-off date: 08-May-2020

[1] 310 patients were randomized but one patient was excluded from analysis as study consent was not signed prior to receiving BAT.

A Cycle was defined as 4 weeks (28 days).

BAT= best available therapy; SR=steroid refractory; aGvHD= acute graft versus host disease; cGvHD= chronic graft versus host disease ORR=overall response rate

2.4.2. Pharmacokinetics

No new drug strength was developed to support this submission.

Bioanalysis

The previously validated method DMB-07.111.2 (Incyte) was used for samples from study 271 /REACH 1. Standard curve and QCs accuracy and precision in the within study validation were within pre-set acceptance criteria. Incurred sample reanalysis was performed for 77 samples (10.3%). All (77/77 or 100%) of the results for INCB018424 agreed within 20% of the original results. The oldest plasma sample was 296 days old prior to analysis, which was within the long term stability of at least 372 days stored at -60 to -80°C.

Two new liquid chromatography tandem mass spectrometry (LC-MS/MS) methods were fully validated for the determination of ruxolitinib in plasma. The method validated at SGS [DMPK R1701012] was employed for the quantitative analysis of ruxolitinib in human plasma samples (with Ethylenediamine-tetraacetic acid (EDTA) as anticoagulant) collected in the Study C2301 as well as in the Study D2301, while the method validated at WuXi AppTec [DMPK R1100270a-02] supported the long term stability of ruxolitinib, but was not used to analyse study samples. The long-term stability period of 813 days (at ≤ -70 °C) covered the maximum length of time from specimen collection to analysis for the two studies.

In both methods, 50 µL K2-EDTA plasma samples were subjected to liquid-liquid extraction, evaporation of the supernatant to dryness, and analysis of the reconstituted sample residue by LC-MS/MS. Validation parameters are summarised in Table 1.

Summary of validation parameters for ruxolitinib methods DMPK R1701012 & R1100270a

Key validation parameters	DMPK R1701012	DMPK R1100270a and amendments 01 and 02
Analyte	Ruxolitinib, INC424-D9 stable labelled internal standard	
Use in current application	Clinical studies C2301, D2301	Long term stability only
Site	SGS (France)	WuXi (China)
Standard curve range (ng/mL)	0.500 ng/mL to 500 ng/mL	0.500 ng/mL to 500 ng/mL
Inter-day accuracy (Bias %)	From −2.67% to −2.00%	From −6.4% to −3.3%
Inter-day precision (CV%)	From 0.86% to 4.78%	From 1.7% to 8.1%
Post-preparative stability in extracts (Auto sampler)	83 hours at 10°C	149 hours at 4°C
Short-term stability in spiked human plasma at RT	24 hours at RT	24 hours at RT
Freeze-thaw stability	3 freeze-thaw cycles at −75°C±10°C	5 freeze-thaw cycles at ≤−15°C and ≤−70°C
Long-term stability when frozen	Not performed*	354 days at ≤−15°C

Long-term stability when deep frozen	250 days at −75°C±10°C	813 days at ≤−70°C

*performed under DMPK R1100270a and its subsequent amendment DMPK R1100270a-01

In the method validation DMPK R1701012, all following parameters were within pre-set acceptance criteria: no interference with fluconazole was demonstrated, no matrix effect (6 individual plasma lots), specificity (8 individual plasma lots), no effect of haemolysis or lipaemia, and dilution linearity upon 10 fold dilution. Certificates of analysis for analytes and internal standard were provided.

The within study validation showed that the standard curve and QCs accuracy and precision were within pre-set acceptance criteria. Incurred sample reanalysis was performed in 47 of study samples from study C2301 and 95.7% of samples met the pre-specified criteria, and in 110 of study samples from study D2301 and 99.1% of samples met the pre-specified criteria.

Study INCB 18424-271 (REACH 1 or 271)

Study 271 was a prospective, open-label, single-cohort, multicenter Phase 2 study of ruxolitinib in combination with corticosteroids for the treatment of steroid-refractory Grades II to IV aGVHD.

Seventy-one participants began treatment at ruxolitinib 5 mg b.i.d; if hematologic parameters were stable and no treatment-related toxicity was observed after the first 3 days of treatment, the dose could be increased to 10 mg b.i.d.

PK samples were taken on Days 1, 7, and 14 (predose and 1 h (±15 min), 2 h (±30 min), and 4-8 h post-dose). PK parameters were calculated from the plasma concentrations of ruxolitinib according to a population PK approach, see below.

Study C2301 (REACH 2)

Two versions of the clinical study report were submitted, the Primary CSR dated 9.3.2020 with data cutoff 25.7.19 with its amendment, dated 12.08.2020 and the Secondary CSR dated 29.07.20 with data cutoff 6.1.2020. Only the primary CSR contained PK data and is presented here.

Study C2301 was a randomized (1:1) Phase III, open label study of ruxolitinib compared to Investigator choice BAT in allogeneic stem cell transplant recipients (adults and adolescents (≥12 years old) with Grade II-IV aGVHD. The assessment of PK of ruxolitinib in SR-aGvHD patients was a secondary endpoint.

Ruxolitinib was administered at a starting dose of 10 mg twice daily (as 2x 5 mg with or without food) until Week 4. Thereafter patients who had complete or partial response continued to receive ruxolitinib up to Week 24. Adolescents received the same ruxolitinib dose as adults. Dose reductions were allowed when given with strong CYP3A4 or dual CYP2C9 and CYP3A4 inhibitors. Dose and frequency of best available therapy (BAT) was based the Investigator's opinion.

The Pharmacokinetic analysis set (PAS) included a total of 200 patients, 152 of whom were patients randomized to ruxolitinib arm, and 48 patients who crossed over to ruxolitinib from BAT. PK parameters from patients with "extensive PK" sampling were available from 22 adult patients and 5 adolescent patients. The sampling was as follows: days 1 and 7 at pre-dose and 0.5 h, 1.0 h, 1.5 h, 2.0 h, 4.0 h, 6.0 h and 9 h (±15 min) post-dose; day 14, 28, 56 and 168: predose and 2 h postdose (± 15 min) for complete and partial responders. Sparse PK sampling was performed for all the other adult patients predose and postdose 2 h (± 15 min) on Days 1, 7, 14, 28, 56, and 168.

PK data was analysed by non-compartmental analysis (NCA), as detailed here, and included in the population PK analysis. The NCA PK parameters were as follows:

After a single dose of 10 mg ruxolitinib, mean plasma concentrations increased rapidly reaching the peak concentration at approximately 1.6 h followed by rapid distribution and elimination phases. On Cycle 1 Day 1, the geometric mean (Geo-mean) peak plasma concentration (Cmax) was 118 (CV 70.4%) ng/mL; AUCinf and AUClast were 470 (CV 71.2%) ng*h/mL and 511 (CV 95.7%) ng*h/mL respectively.

After multiple dosing at Day 7, the accumulation ratio was low (1.2), indicating minimal accumulation of ruxolitinib with continuous dosing. After 7 days of continuous dosing, geometric mean Cmax was 137.1 ng/mL and AUCtau was 713.4 ng.h/mL.

Oral plasma clearance at steady state (CLss/F) of ruxolitinib after oral dosing at 10 mg twice daily was estimated as 23.3 L/h, which was comparable to the CL/F obtained from AUCinf after a single 10 mg oral dose. This suggested that clearance is not expected to change with time. Similarly, geometric mean terminal T1/2 appeared to be independent of time and ranged from 1.9-2.0 h after both single and repeated doses.

Most of the patients in the ruxolitinib arm (134/152) received concomitant CYP3A4 inhibitors during the study. The most frequently administered CYP3A4 inhibitors were cyclosporin (61.2%), posaconazole (48.0%) and voriconazole (20.4%) which are known to be moderate or strong inhibitor of CYP3A4. Though a large fraction of the ruxolitinib-treated population received CYP3A4 inhibitors, PK parameters on Day 7 could be estimated only for 22 patients from whom extensive PK samples were collected. Of these 22 patients, only 4 patients were without a strong or moderate concomitant CYP3A4 inhibitor and the PK variability in these 4 patients was high.

PK parameters at Day 1 in adolescent patients could be computed for only 5 patients, however, within the ambit of the available data, the exposure in adolescent patients was within the range observed in adult patients (Table below).

Individual observed PK parameters in adolescent patients with acute GvHD after single and repeated dosing

Patient	Age (years)	BSA (m^1)	BW (kg)	Cmax (ng/mL) Day 1	Cmax (ng/mL) Day 7	AUClast* (ng*h/mL) Day 1	AUClast* (ng*h/mL) Day 7
patient 1	15	1.43	45	128	NA	366	NA
patient 2	13	1.45	46	111	146	489	733
patient 3	16	2.16	97	156	NA	963	NA
patient 4	12	1.03	28.5	357	165	809	555
patient 5	13	1.19	32.9	149	363	710	2010
Min-Max range observed in adolescents				111-357	146-363	366-963	555-2010
Min-Max range observed in adults				20.8-580	33.8-744	63.8-4484	116-3064

*PK samples were collected up to approximately 9 hours post-dose

Study D2301 (REACH 3)

Study D2301 was a randomized, open-label, multi-center study of ruxolitinib vs. BAT in patients (adults and adolescents (≥ 12 years old) with moderate or severe cGvHD after allogeneic stem cell transplantation. The assessment of PK of ruxolitinib in SR-cGvHD patients was a secondary endpoint. Ruxolitinib vs. BAT was added to the patient' s immunosuppressive regimen corticosteroids ± calcineurin inhibitor (CNI). Dose and frequency of BAT was based the Investigator's opinion.

Ruxolitinib was administered at a dose of 10 mg b.i.d daily, (as 2x 5 mg with or without food) for a minimum of 6 (28 day) cycles until cycle 7 Day 1. Thereafter patients who had complete or partial response continued to receive ruxolitinib up to Cycle 39. Dose modifications (reductions and/or interruptions) and delays were permitted to address safety concerns related to haematological (primarily, cytopenias) or non-haematological reasons, or to avoid under- or over-exposure of ruxolitinib when combined with CYP450 modulators. Adolescents received the same ruxolitinib dose as adults.

The Pharmacokinetic analysis set (PAS) included a total of 221 patients, 164 of whom were patients randomized to ruxolitinib arm, and 57 patients who crossed over to ruxolitinib from BAT. PK parameters from patients with "extensive PK" sampling were available from 17 adult patients and 4 adolescent patients who received a dose of 10 mg ruxolitinib twice daily. PK parameters at Day 1 in adolescent patients could be computed for only 2 or 3 available samples for one of the adolescents.

For the extensive PK dataset, on days 1 and 15 of Cycle 1 samples were taken pre-dose (0 h), post-dose (0.5 h, 1 h, 1.5 h (±15 min), 4 h, 6 h, 9 h (±1 hr); at Cycle 2, 7 and 39, one pre-dose and one-post dose 1.5 hr (±15 min). Sparse PK sampling was available for all the other patients: predose and postdose 1.5 h (± 15 min) on Day 1 of Cycle 1, Day 15 of Cycle 1, and Day 1 of Cycle 2, 7 and 39. Sparse PK samples for cross-over patients from BAT to ruxolitinib were taken only at the time of a scheduled visit.

PK data was analysed by non-compartmental analysis, as detailed here, and included in the population PK analysis.

After a single dose of 10 mg ruxolitinib, median Tmax was 0.833 h; geo mean Cmax was 167 (39.3%) ng/mL; AUCinf and AUClast were 642 (32.7%) ng.h/mL and 636 (40.8%) ng.h/mL respectively; elimination half-life (T1/2) was 2.40h in patients receiving ruxolitinib at a dose of 10 mg.

After multiple dosing at Day 15, ruxolitinib PK profile was similar to that at Day 1. A slightly higher peak plasma concentration level was observed in the Day 15 profile; however, large variability was noted. On Day 15, geo mean AUCtau was 656 (20.4%) ng.h/mL; AUClast was 945 (56.1%) ng.h/mL and Cmax was 215 (48.8%) ng/mL in patients receiving ruxolitinib at a dose of 10 mg b.i.d.

Oral plasma clearance at steady state (CLss/F) of ruxolitinib after oral dosing at 10 mg twice daily was estimated as 15.2 L/h. The geometric mean terminal half-life (T1/2) appeared to be independent of time and ranged from 2.3 to 2.4 h.

Within the ambit of the available data, the exposure in adolescent patients was within the range observed in adult patients (Table below).

Individual observed PK parameters in adolescent patients with chronic GvHD after single and repeated dosing

Patient	Age (years)	BSA (M^2)	BW (kg)	Cmax (ng/mL) Day 1	Cmax (ng/mL) Day 15	AUClast* (ng*h/mL) Day 1	AUClast* (ng*h/mL) Day 15
patient 1	13	1.37	45	249	253	773	711
Patient 2	15	1.2	32.5	152	286	595	849

Patient 3	16	2.31	107	190	274	783	1120
Min-Max range observed in adolescents				152-249	253-286	595-783	711-1120
Min-Max range observed in adults				67.8-312	95.5-451	313-1400	450-2720

*PK samples were collected up to approximately 9 hours post-dose

Population Pharmacokinetic analysis

The applicant has submitted three model based pharmacokinetic analysis reports.

1. Report DMPK R1800133 (Projected Pharmacokinetic Parameters and Anticipated Therapeutic Dose of Ruxolitinib (INC424) in Paediatric population) describes dose selection in children from 0 years and is out of the scope for this application. Therefore, it is not reviewed.

2. Report DMB-18.34.1 (Population Pharmacokinetic Analysis of the JAK Inhibitor Ruxolitinib Tablets in Combination with Corticosteroids for the Treatment of Steroid-Refractory Acute Graft-Versus-Host Disease INCB 18424-271) only includes GvHD patient data from Reach 1 (phase 2 study). The report below is updated with additional phase 3 data, has similar objectives, and includes a re-assessment of the final model presented in this report. Therefore, this model is not assessed in dept or described in this assessment.

3. Report named Population pharmacokinetics of ruxolitinib in steroidrefractory (SR) acute and chronic (a+c) Graft versus Host Disease (GvHD) adult and adolescent patients Modeling Report (Final, 27 Nov 2020) includes GvHD patients' data from Reach 1, 2 and 3 and an update of the model from Report DMB-18-34-1. This report is considered to include the pivotal analysis of ruxolitinib in the GvHD population and is described below.

Methods

The objective was to characterize PK in acute and chronic GvHD patients and investigate the covariates in this population which explain the inter-patient variability.

Software

The analysis is performed using the Monolix software system, Monolix 2019R2 (Lixoft, Paris, France) utilizing the DaVinci high performance computing environment accessed from GPSII. All model building was performed using Stochastic Approximation Expectation Maximization (SAEM) method).

Data

The final model included data from the studies Reach 1, 2 and 3, added during model development in a staggered manner. 11.6% of data were censored (below quantification limit).

Model

The model was a two-compartment model with first-order absorption and linear elimination, with a lag time, and an exponential residual error structure (additional error model, with log-normal IIV distributions). Inter-individual variability (IIV) was estimated for all parameters apart from lag time and inter-compartmental clearance. Administration of CYP3A4-inhibitors (moderate and potent vs weak and none), aGvHD disease grade at each assessment (MAGIC score, <4 vs. 4), and liver involvement in aGvHD at each assessment (liver score >=1, Y/N), all included on apparent clearance, and aGvHD disease grade at each assessment (MAGIC score <=1 vs. >1) on the absorption rate. A covariance term was added for the IIV on apparent central volume and clearance.

Components of this model were reassessed to determine whether simpler or alternative models could better describe the PK data in aGvHD patients. In particular, according to [PK Report DMB-18 341], high IIV was observed on V2 and ka, which was attributed to the composition of the data used for model building, notably the limited dense PK sampling for REACH 1 (and also 1 of the MF studies).

The original model incorporates an additive error model, with log-normal distributions implemented on the IIV terms. A normal distribution will be implemented in place of a log-normal distribution.

The model was reassessed first by adding data from study Reach 2, and subsequently reassessed again adding data from study Reach 3. All comparisons will be made using the BIC, since models will not be nested and therefore a likelihood ratio test cannot be performed. A reduction of >3 BIC units constitutes a significant model improvement.

Relevant covariates will be assessed with the aGvHD final base model, along with additional potential candidates that may be expected to impact ruxolitinib PK. Forward selection will be implemented and correlated covariates will be tested separately and the more appropriate covariate will then be included individually. Eighty percent of the datasets were used for covariate model building, with the remaining 20% used for validation. The data was split with a random Bernoulli variable. Patients with missing/unknown levels of a covariate of interest (e.g. Race) were also excluded from the model building dataset and added to the external validation dataset provided they were not a significant covariate in the final model. All data (with valid covariate information as per the final model) was used for simulations.

Final Model

The final model is a 2-compartment, 1st order absorption model, with a lag time and linear elimination. Additional IIV was placed on CL/F and Vc/F.

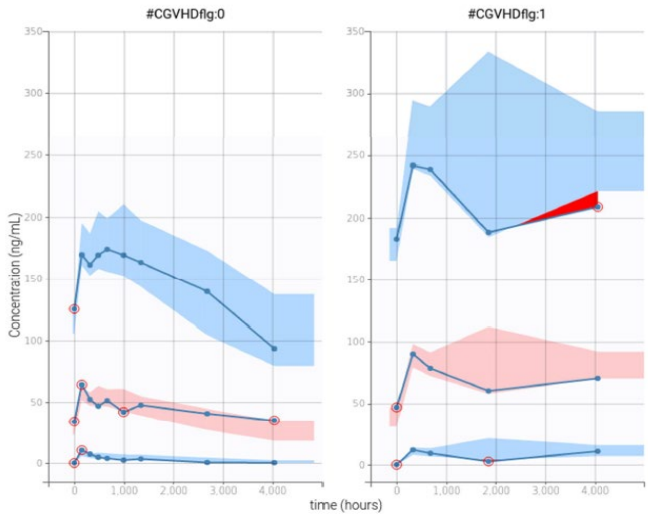
Median scaled BSA had a significant impact on Vc/F and Cl/F, with an increase in Vc/F observed with increasing BSA (as expected), and decreased Cl/F with increasing BSA. Lower GI involvement to the GvHD at baseline had a significant impact on primarily on the peripheral volume, but also on the absorption rate. The presence of lower GI involvement of the disease at baseline indicated a lower absorption rate and an increase in the peripheral volume. The reduction in absorption rate is expected, since ruxolitinib is orally administered, and therefore any indication of GI disease would logically impact the absorption of the drug. The indication of chronic GvHD (compared to acute GvHD) has a significant impact on Ka, Cl/F and Vc/F, resulting in an increased Ka, decreased Cl/F and decreased Vc/F. All of these impacts result in overall higher exposures in cGvHD patients compared to aGvHD patients, as previously indicated by the initial VPC (for the final aGvHD model on the cGvHD data).

The final parameter estimates, after running the model with an increased number of simulations, are given below and the models ability to predict the observed data, stratified on acute and chronic GvHD is presented in Figure 1, and BSA <1.5 mf vs >=1.5 m2 is presented in Figure 2. The impact of BSA and other covariates is graphically presented in Figure 3 and Figure 4.

Final parameter estimates

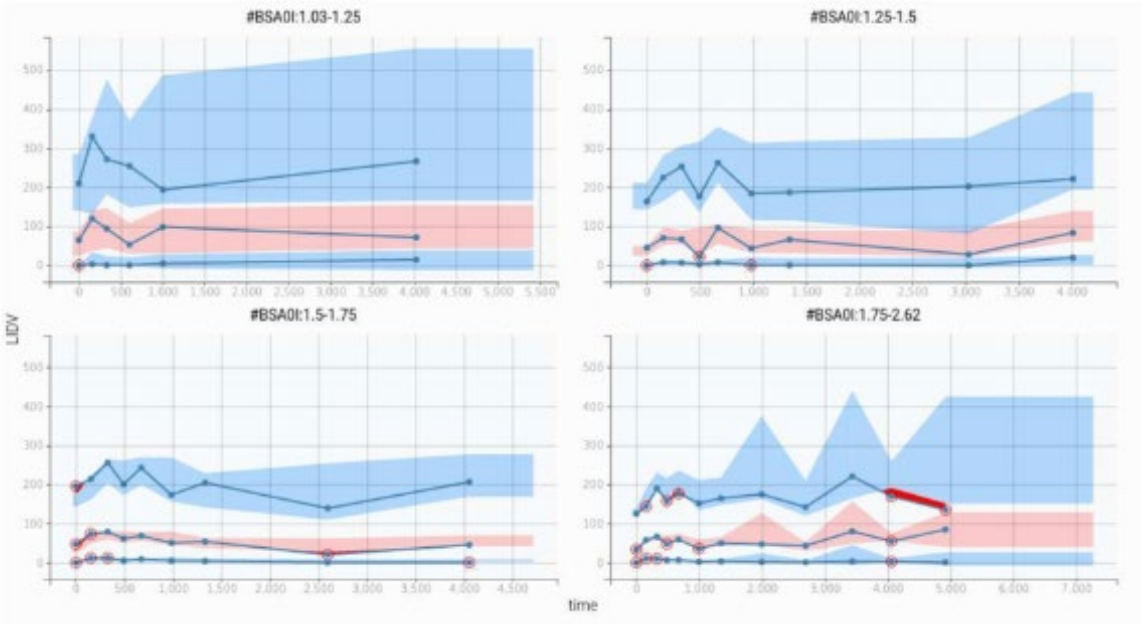
Parameter	Population Estimate			IIV (Omega)		
	Mean	SE	RSE (%)	Mean (SD)	SE	RSE (%)
Ka (h^{-1})	3.35	0.951	28.4	1.16	0.128	11
Involvement of lower GI at baseline on Ka	-1.06	0.291	27.5	-		
cGvHD on Ka	0.636	0.34	53.4			
Tlag (h)	0.195	0.0376	19.3	0.858	0.126	14.7
Vc/F (L) for a patient with BSA=1.83	17.5	1.16	6.59	0.145	0.00721	4.99
BSA/1.83 on Vc/F	1.35	0.0648	4.8			
cGvHD on Vc/F	-0.103	0.019	18.4			
Cl/F (L/h) for a patient with BSA=1.83	6.57	1.47	22.4	0.508	0.0247	4.86
BSA/1.83 on Cl/F	0.46	0.219	47.7	-		
cGvHD on Cl/F	-0.286	0.0641	22.4			
Vp/F (L)	4.89	4	81.8	4.99	0.487	9.75
Involvement of lower GI at baseline on Vp/F	2.9	0.954	32.8	-		
Q (L/h)	13.8	0.586	4.25	NA		
Residual variability (CV)	0.487	0.00771	1.58	NA	NA	NA
Correlation between Cl/F, Vc/F	0.999	0.000613	0.0613	NA	NA	NA

Figure 1. Visual Predictive Check of patients with aGvHD (left plot, from REACH 1 or 2) and cGvHD (right plot, from REACH 3)



Solid lines display observed 10th, 50th, 90th percentiles
Blue/Pink regions show 90% prediction interval around the percentiles
Circled dots show observed percentiles outside of the prediction interval

Figure 2. Visual Predictive Check of patients stratified on BSA (1-1.25m2 (n=9), 1.25-1.5m2 (n=47), 1.5-1.75m2 (n=133), >1.75m2 (n=293))



Solid lines display observed 10th, 50th, 90th percentiles
Blue/Pink regions show 90% prediction interval around the percentiles
Circled dots show observed percentiles outside of the prediction interval

Figure 3. Simulated AUC 0-12 (A) and Cmax (B) on day 1 by indication, across the BSA range

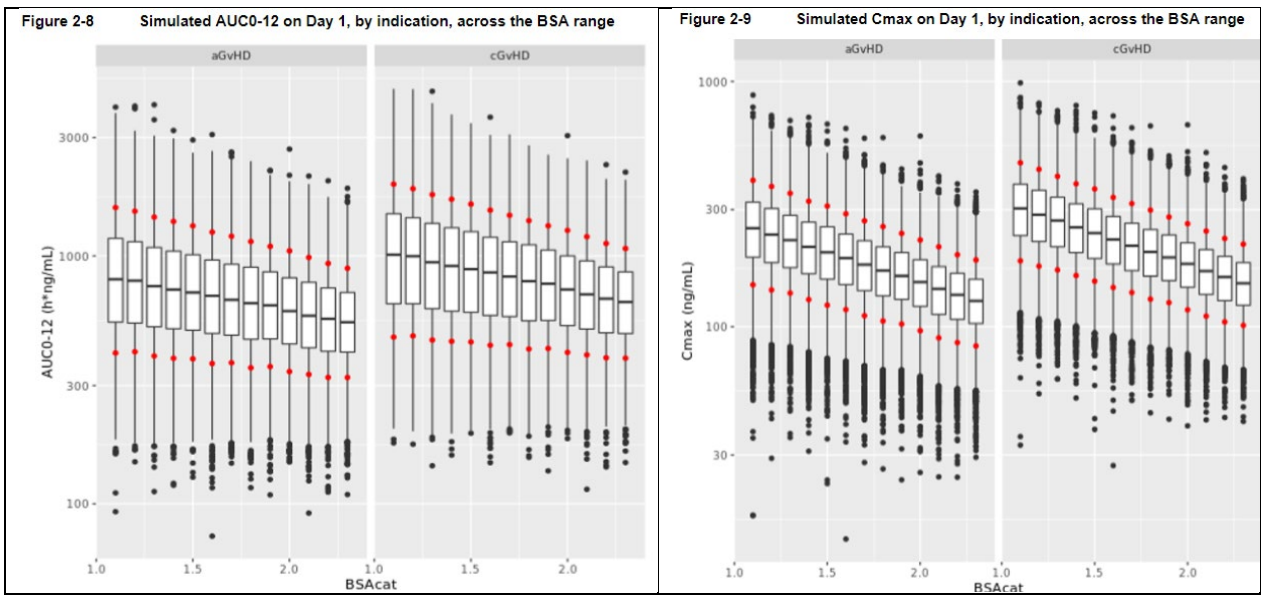
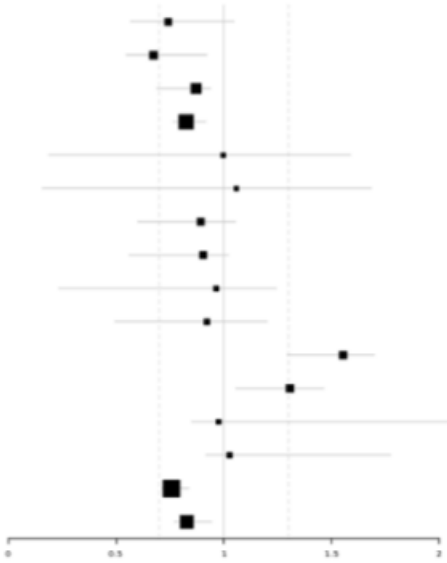


Figure 4. Forest Plot of covariate effects on PK parameters, including BSA

Parameter	Ref.: LGI=N Chronic GvHD BSA=1.79	% diff
C12h D1	aGvHD vs. cGvHD	-25.86
Ctrough D7	aGvHD vs. cGvHD	-32.65
Cmax	aGvHD vs. cGvHD	-12.88
AUC0-12	aGvHD vs. cGvHD	-17.52
C12h D1	LGI=Y vs. LGI=N	-0.32
Ctrough D7	LGI=Y vs. LGI=N	5.84
Cmax	LGI=Y vs. LGI=N	-10.75
AUC0-12	LGI=Y vs. LGI=N	-9.57
C12h D1	BSA=1.06 vs. BSA=1.79	-3.58
Ctrough D7	BSA=1.06 vs. BSA=1.79	-7.9
Cmax	BSA=1.06 vs. BSA=1.79	55.35
AUC0-12	BSA=1.06 vs. BSA=1.79	30.72
C12h D1	BSA=2.23 vs. BSA=1.79	-2.4
Ctrough D7	BSA=2.23 vs. BSA=1.79	2.66
Cmax	BSA=2.23 vs. BSA=1.79	-24.37
AUC0-12	BSA=2.23 vs. BSA=1.79	-17.25



2.4.3. PK/PD modelling

Exposure-Response

Data from the two phase 3 studies (C2301 for aGvHD and D2301 for cGvHD) were used and analysed separately. Time-averaged AUC0-12h was the exposure metrics that was used. Cmax was not considered for safety endpoints, since the safety endpoints selected were chosen based on underlying disease and systemic treatments. Several demographic and disease related covariates were investigated in the analysis. If there were more than 10% missing values for a covariate then the covariate was not included in the covariate selection.

Overall response rate (ORR), durable response rate (DRR) and modified Lee symptoms score at Cycle 7 Day 1 were analysed with logistic regression models. Overall Survival (OS), duration of response (DoR) and failure free survival (FFS) were analysed with Cox regression models. The PK-efficacy set included 150 patients each from Study C2301 (up to the data cut-off date of 06-Jan-2020) and Study D2301 (up to the data cut-off date of 15-May-2020).

Only adverse events of special interest (AESI) that were reported in adequate numbers (at least 10% and no more than 90% of the patients in the PK Safety Set of each AESI) of patients were analysed. If there was a considerable portion of patients with several occurrences of the same AESI (e.g. if the median number of AESI per patient is 2 or more), this AESI in question was analysed with a count regression model. The PK-safety set included 194 patients from Study C2301 (up to the data cut-off date of 06- Jan-2020) and 150 patients from Study D2301 (up to the data cut-off date of 15-May-2020)

Results

There was an absence of a strong relationship between the exposure metric (time-averaged AUC0-12) and the efficacy and safety endpoints for patients with aGvHD and cGvHD. This may be due to the limited exposure range observed with a dosing regimen of 10mg bid.

Population pharmacokinetic/pharmacodynamic relationship of ruxolitinib and platelet count in GvHD patients

There were two main objectives with the below presented analysis

1. To characterize exposure–response relationship of ruxolitinib and longitudinal platelet count, in the context of thrombocytopenia, in GvHD (acute+chronic) patients and investigate the covariates in this population which could explain a part of the inter-patient variability

2. Quantify and qualify the impact of ruxolitinib dose on the ruxolitinib induced thrombocytopenia in acute and chronic GvHD population through simulations of patient platelet profiles, for fixed dosing regimens of 5mg bid and 10mg bid, accounting for predictors of platelets counts identified during model building.

Methods

Average predicted concentration between platelet measurements was used instead of observed concentrations (derived from the final pop PK model).

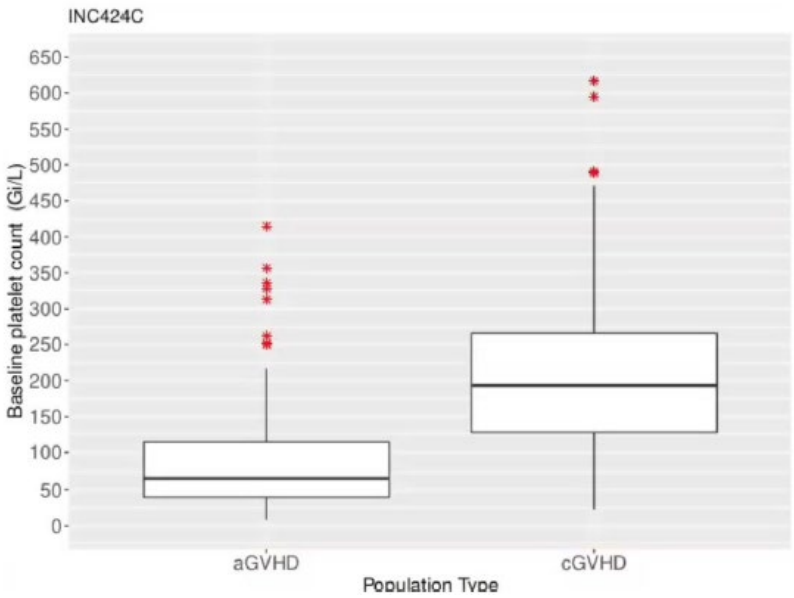
A previously developed base semi-mechanistic life span model was modified in order for model to be able to capture the typically very low (<50 Gi/L) baseline platelet counts seen in acute GvHD patient. A small set of covariates were tested in the final acute and chronic GvHD model. The number of patients, platelet and transfusion observations per study in the corresponding modelling datasets is listed below. The modelling datasets included platelet and transfusion observations for the first 168 days of treatment.

Number if platelet observations and platelet transfusions by study in the modelling datasets

Indication/Modeling dataset	Number of patients	Number of platelet observations	Number of patients with platelet transfusions	Number of transfusions
acute, test data	39	292	22	212
acute, training data	153	1448	84	862
chronic, test data	45	393	3	6
chronic, training data	172	1416	12	67

Distinct differences between the two studies were observed with baseline platelet count, as well as overall profiles whilst receiving ruxolitinib treatment (Figure 5). This is largely due to a longer time since stem cell transplant in the cGvHD indication, by definition, >100 days. For the aGvHD patients (from study C2301), there is limited trend. This is in general due to the fact that all patients are <100 days since transplant, and therefore have not fully achieved platelet recovery.

Figure 5. Boxplot of baseline platelet counts stratified by study



Given the differences observed between the two studies, and in particular, the fact that study D2301 (cGvHD patients) have generally normal platelet levels, the analysis moving forward will be conducted only on study C2301 (aGvHD patients).

Final model

The final base model therefore incorporated a chain of 5 compartments, including 1 precursor compartment with proliferative cells, such as stem cells and other progenitor cells, three transit compartments with maturating cells, and one compartment for circulating platelets. A feedback mechanism was incorporated to reproduce proliferation regulation based on ratio between normal baseline value and actual platelet count. In the acute GvHD population most patients start with extremely low platelet counts due to the short time since transplantation, and the inclusion of time-varying baseline platelet count is needed to capture these observations. Plasma concentrations of ruxolitinib (averaged between platelet measurements) were assumed to inhibit the proliferation rate by a linear slope model. Platelet transfusion events were incorporated in the model to account for resulting quick elevations in platelet counts. Only the covariate "effect of time since the engraftment with the respect to the first dose" (TTRANSDOS1) on projected platelet count at engraftment time (PLTT) was carried into the final model.

The parameter estimates for the final model are shown in Table below. The final fit of the model is displayed in Figure 6, with a further validation component to fit the model to a separate test dataset (20% of the original dataset) and is displayed in Figure 7. The VPC for the covariate model improves the fit towards the end of the treatment period, suggesting that by adjusting for time since transplant (relative to first dose) on the platelet count, allows a better description of the recovery phase.

Parameter estimates for the final acute semi-mechanistic platelet model

Parameter (units)	Population Estimate			Random effects		
	Mean	SE	RSE (%)	Mean (SD)	SE	RSE (%)
ktr12	0.257	0.0207	8.06	0.1 FIX	-	-
α $(10^9/L)$	40.6	3.79	9.34	0.1 FIX	-	-
iLAG (hr)	18.9	0.666	3.52	-	-	-
MMT (hr)	191	22.1	11.6	0.631	0.0817	12.9
γ (-)	0.29	0.0223	7.66	-	-	-
PLTT0$(10^9/L)$	79.2	7.67	9.68	0.583	0.0429	7.35
PLTT0~TTRANDOS 1	-0.000132	3.86e-5	29.3			
SLOPE	4.1e-5	4.94e-6	12	0.764	0.105	13.8
w $(10^9/L)$	142	12.2	8.64	0.3 FIX	-	-
t50(days)	260	20.1	7.74	0.553	0.0808	14.6
j (-)	10 FIX				-	-
Corr(PLTT,MMT)	0.583	0.118	20.3			
Residual error model						
a	5.38	0.952	17.7	NA	NA	NA
b	0.332	0.0125	3.76	NA	NA	NA

Ktr12=kout: first-order platelet count removal rate constant, α: bioavailability of platelets for transfusion

iLAG: platelet transfusion lag time; MMT: Mean transit time of platelet life cycle

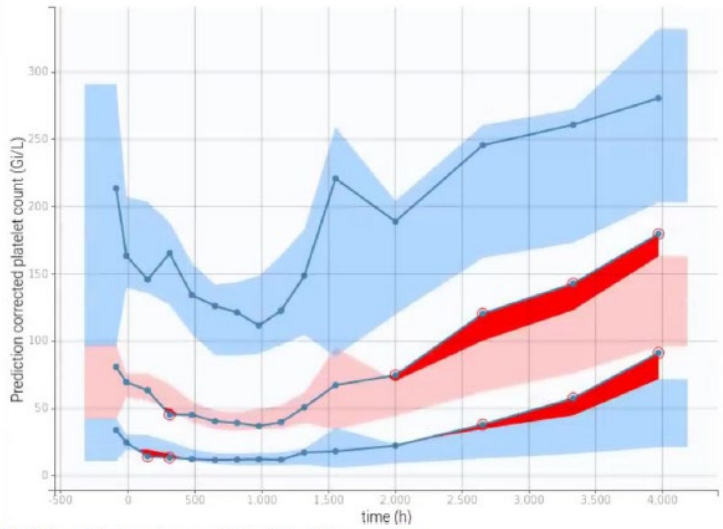
γ: coefficient associated with feedback mechanism

PLTT0: projected platelet count at engraftment time; SLOPE: proportional drug effect on precursor cells; w: the maximum possible increase of baseline platelet count since engraftment

t50: time since engraftment with half of recovery

j: Hill coefficient

Figure 6. Visual predictive check of the final aGvHD platelet model in the training dataset

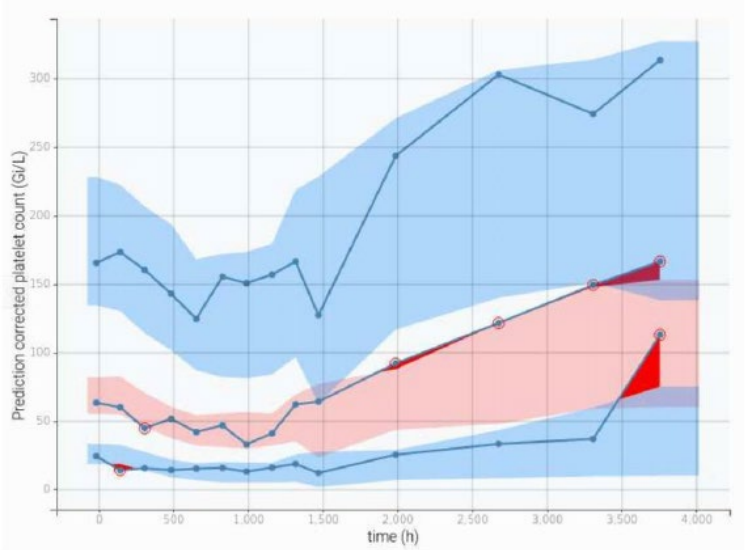


Solid lines display observed 10th, 50th, 90th percentiles

Blue/Pink regions show 90% prediction interval around the percentiles

Circled dots show observed percentiles outside of the prediction interval

Figure 7. Visual predictive check of the external validation dataset



Solid lines display observed 10th, 50th, 90th percentiles
Blue/Pink regions show 90% prediction interval around the percentiles
Circled dots show observed percentiles outside of the prediction interval

The model well captures the initial reduction of platelet count, initiated from a very low baseline, and gradual recovery of platelet function (recovery is defined as >100,000/mm3) which is observed while still receiving ruxolitinib (within 4032 hours, which corresponds to Day 168 of study). Extrapolation of this model to high platelet count is somewhat optimistic in the early phase. The full recovery generally occurs before 100 days (by 2400 hours, 3 cycles) which is during the taper period. The results suggest that for aGvHD patients, there is an initial reduction of platelet count with initiation of ruxolitinib, but this can generally be managed well by short term dose reductions/interruptions and platelet transfusions as per protocol.

2.4.4. Discussion on clinical pharmacology

The bioanalytical method used for study 271 / REACH1 was previously assessed in procedure EMEA/H/C/002464/II/16. Within study validation for study 271 is adequate. Both methods DMPK R1701012 & R1100270a have demonstrated acceptable sensitivity, specificity, selectivity and stability for the quantification of ruxolitinib. Within study validation was also adequate. Methods DMPK R1701012 & R1100270a are not cross-validated. This is acceptable since method R1100270a is only used to support the long term stability of samples and is itself adequately validated.

Single and repeated PK data following a 10 mg ruxolitinib were obtained in acute GvHD and chronic GvHD subjects. In both groups, ruxolitinib was absorbed rapidly, with a tmax of approximately 1.5 hour. After multiple dosing, the accumulation ratio was low (1.2), indicating minimal accumulation of ruxolitinib with continuous dosing. Terminal t1/2 was approximately 1.9-2.0h in aGvHD subjects and 2.3-2.4h in cGvHD subjects. Further, ruxolitinib PK appeared to be independent of time. Clearance was 10.4 l/h in patients with acute GvHD and 7.8 l/h in patients with chronic GvHD, with a 49% inter subject variability. No relationship was apparent between oral clearance and gender, patient age or race, based on a population pharmacokinetic evaluation in GvHD patients.

Special populations & interactions

Compared to the dedicated studies, no mechanistic or physiological difference is expected in the GvHD population for patients that have hepatic impairment (HI) or when combined with CYP3A4 inhibitors that would result in a different outcome with respect to PK. Therefore, a similar magnitude of

interaction is expected in GvHD patients and this is aligned to the recommendation for CYP3A4 inhibition and for hepatic impairment not due to GvHD with prior recommendation.

Dual CYP3A4/CYP2C9 inhibitors. In line with the MF and PV indications and partly based on the popPK analysis where visual predictive checks suggested that administration of fluconazole in patients increased ruxolitinib concentration levels approximately 2-fold, the 50% dose reduction in patients with GvHD when co-administered with dual inhibitors of CYP2C9 and CYP3A4 enzymes is considered acceptable.

Hepatic impairment. Based on the observed and predicted plasma exposures across all hepatic impairment categories (total bilirubin and AST levels) in the PopPK analyses no dose adjustment appears necessary in GvHD patients with mild, moderate, or severe hepatic impairment. Nevertheless, considering the numbers of aGvHD and cGvHD patients included in each hepatic impairment subgroup (as classified by total bilirubin and AST levels) in the popPK analysis and the data collected, the proposed conservative approach and the advice that the starting dose of ruxolitinib in patients with liver impairment not related to GvHD should be reduced by 50%, in line with the current advice in MF and PV patients, is agreed.

Renal impairment. In the popPK study, GvHD patients with moderate renal impairment had very similar observed and predicted plasma exposures to those who had no renal impairment at baseline. This provides support that the current posology advice for MF and PV patients with mild or moderate renal impairment patients, i.e. no dose adjustment is recommended, is also valid for GvHD patients belonging to the same special subpopulation.

Considering the proposed recommendation on dosing of GvHD subjects with ESRD and in absence of further modelling and simulation data the SmPC has been updated including the proposed text that no data are available for ESRD.

The pharmacokinetics profile observed in adolescent patients with acute or chronic GvHD was comparable to the overall patient population

Population PK

The main objective of the model-based population PK analysis was to assess relevant covariates to explain the exposure in patients with acute and chronic GvHD and have a model adequate for simulation of post-hoc PK parameters to be used in exposure-response analysis. Adolescent subjects are expected to have a lower, or much lower, BSA than the average adult. In addition, the GvHD population is more sensitive to haematological side effects. It is of importance that the starting dose in the subjects with lowest BSA is appropriate. The presented VPCs indicate that the model can capture the general trend in the analysed population (no major deviations between observed and predicted data points), and the goodness-of-fit plots to not indicate any obvious major trends of model misspecification. There is a concern regarding the potential overparameterization of the final model. However, the VPCs adequately predict the data with respect to BSA and acute and chronic GvHD population and the final model parameters seem reasonable (i.e. not large RSEs or unreasonable parameter estimates). Therefore, given the application of the model (inform the suitability of the dose in subjects with low BSA) and that there is sufficient clinical efficacy and safety data, a new model is not requested. The safety and efficacy data are mainly represented by subjects with a BSA >1.5m^2. (BSA 1.03-1.25m^2 n=9; BSA 1.25-1.5 m^2, n=47; BSA 1.5-1.75 m^2, n=133; BSA >1.75 m^2 n=293). All subjects included in the study received the same starting dose, which was adjusted based on safety and efficacy. BSA is the covariate that has the highest impact on exposure of ruxolitinib, however there is a considerable overlap of exposures. The impact of BSA on ruxolitinib exposure is presented in the SmPC section 5.2.

PK/PD modelling

The exposure-response analysis was sufficiently explored, however the results from the analysis are inconclusive and should be interpreted with caution. This is because all subjects included in the exposure-response analysis received 10 mg b.i.d. which results in a limited exposure range. There is sufficient clinical patient efficacy and safety data to overcome the inconclusive data from Modelling. (see Clinical Efficacy and Clinical Safety sections).

The applicant developed a model to characterize the relationship of ruxolitinib and longitudinal platelet count, in the context of thrombocytopenia, in acute GvHD patients. Only data from the phase 3 C2301 study (aGvHD) study was used in the model development. There is limited information on impact of starting dose and subsequently exposure, because all patients had a starting dose of 10 mg b.i.d. Therefore, information available is too limited to answer the second objective (i.e. impact of starting dose on thrombocytopenia) and results of the simulations should be interpreted with caution.

PD. For the pivotal clinical Phase 3 trials, PD related endpoints were defined to explore cytokines and GvHD biomarkers were planned to be collected and included as exploratory endpoint in both pivotal trials. Results of these exploratory endpoints have not been presented. No conclusion on predictive values of the biomarkers from the selected panels can be drawn at this point. Evaluations for the association of selected marker with response are continuing as the data matures and are expected to be reported when available.

2.4.5. Conclusions on clinical pharmacology

Overall, ruxolitinib PK in GvHD subjects is clarified to a sufficient extent.

Population PK model adequately characterize the PK of ruxolitinib in the GvHD population. Body surface area was the covariate with highest impact on ruxolitinib exposure. Subjects with lower BSA have a higher exposure, which is reflected in the SmPC. The results from the exposure-response analysis are inconclusive and should be interpreted with caution because all subjects included received 10 mg b.i.d. which results in a limited exposure range. The applicant developed a model to characterize the relationship of ruxolitinib and longitudinal platelet count in acute GvHD patients using only data from the phase 3 C2301 study (aGvHD), however, there is no information on impact of a different starting dose.

2.5. Clinical efficacy

2.5.1. Dose-selection rationale

No formal dose finding study has been presented. Ruxolitinib dose selection was based on the published data on preliminary efficacy and safety generated in patients with SR-GvHD and from the supportive study INCB 18424-271 (A Single-Cohort, Phase 2 Study of ruxolitinib in combination with corticosteroids for the treatment of steroid-refractory acute Graft-Versus-Host Disease [REACH-1])(Zeiser et al 2015). The dose administered in this study was lower than that generally administered in myelofibrosis patients (i.e. 15-20 mg bid), due to that alloSCT patients with SR-aGvHD would routinely be treated with concurrent CNI and azole prophylaxis that can inhibit the metabolism (via CYP3A4) of ruxolitinib, potentially increasing its exposure. The usage of same dose of ruxolitinib 10 mg b.i.d. in adolescents in Study C2301 and D2301 was supported by the published literature showing similar toxicity profiles, maximum tolerated doses, and ruxolitinib PK parameters in both adolescents

and adults (Loh et al 2015). The dose selection seems adequate, based on preclinical and clinical publications and Pk analysis.

2.5.2. Main studies

There are two pivotal Phase III Studies (CINC424C2301 and CINC424D2301, hereafter referred to as study C2301 and study D2301 or REACH2 and REACH3) presented in the present dossier, designed as randomized, multi-centre, open-label studies to investigate the efficacy and safety of ruxolitinib (RUX) compared to Investigator's choice of best available therapy (BAT) in patients with SR-aGvHD and SR-cGvHD, respectively.

2.5.3. Steroid refractory acute GvHD

Title of Study: A phase III randomized open-label multi-centre study of ruxolitinib versus best available therapy in patients with corticosteroid-refractory acute graft vs. host disease after allogeneic stem cell transplantation (INC424C2301).

Name of product: ruxolitinib/INC424

Study number: CINC424C2301 (EudraCT number: 2016-002584-33). Study identifier: **REACH2**
Participating countries/regions and (number of sites): EU (65, incl. UK and Norway), Australia (4), Canada (8), Hong Kong (1), Israel (4), Japan (14), Republic Of Korea (2), Russian Federation (1), Saudi Arabia (1), Taiwan (1), Turkey (4).

Study period

Study initiation date: 10-Mar-2017 (first patient first visit)
Data cut-off date: 25-Jul-2019 (data cut-off date for primary analysis) and 06-Jan-2020 (data cut-off date for second analysis); study is ongoing. The final analysis will occur once all patients have completed the Long-Term Follow-Up period to 24 months. All available data from all patients up to EOS, inclusive of OS, will be reported in a final CSR.

Phase of development (phase of this clinical study): Phase III

Methods

Study participants

Inclusion criteria

Patients eligible for inclusion in the study had to meet all of the following criteria:

1. Written screening informed consent and/or assent from the patient, parent, or guardian at the time of Screening, i.e. at the time of aGvHD Grade II-IV diagnosis.
2. Written study informed consent and/or assent from the patient, parent, or guardian once SRaGvHD was confirmed.
3. Male or female patients aged 12 or older at the time of screening informed consent.
4. Able to swallow tablets.
5. Had undergone alloSCT from any donor source (matched unrelated donor, sibling, haploidentical) using bone marrow, peripheral blood stem cells, or cord blood. Recipients of non-myeloablative, myeloablative, and reduced intensity conditioning were eligible.

6. Clinically diagnosed Grades II to IV acute GvHD as per standard criteria occurring after alloSCT requiring systemic immune suppressive therapy. Biopsy of involved organs with aGvHD was encouraged but not required for study screening.
7. Evident myeloid and platelet engraftment (confirmed within 48h prior to study treatment start):
    - Absolute neutrophil count (ANC) > 1000/mm3 and
    - Platelets ≥ 20,000/ mm3

Note: Use of growth factor supplementation and transfusion support was allowed.

8. Confirmed diagnosis of SR-aGvHD

Key exclusion criteria: Patients eligible for this study should not have met any of the following key criteria:

- Received more than one systemic treatment for SR-aGvHD.
- Clinical presentation resembling de novo chronic GvHD or GvHD overlap syndrome with both acute and chronic GvHD features (as defined by Jagasia et al 2015).
- Failed prior alloSCT within the past 6 months.
- Presented with active uncontrolled infection requiring treatment.
- Presented with relapsed primary malignancy, or patients who were treated for relapse after the alloSCT was performed, or who may require rapid immune suppression withdrawal as pre-emergent treatment of early malignancy relapse.
- SR-aGvHD occurring after non-scheduled donor lymphocyte infusion (DLI) administered for pre-emptive treatment of malignancy recurrence.
- Significant respiratory disease including patients who were on mechanical ventilation or who have resting O2 saturation <90% by pulse-oximetry
- Severely impaired renal function defined by serum creatinine >2 mg/dL (>176.8 μmol/L), renal dialysis requirement, or have estimated creatinine clearance <30 mL/min measured or calculated by Cockroft Gault equation.
- Clinically significant or uncontrolled cardiac disease including any of the following:
    - Acute myocardial infarction within 6 months from Day 1 of study treatment administration
    - Uncontrolled hypertension
    - New York Heart Association Class III or IV congestive heart failure
    - Unstable angina within last 6 months from screening
    - Clinically significant (symptomatic) cardiac arrhythmias (e.g., sustained ventricular tachycardia, and clinically significant second or third degree atrio-ventricular block without a pacemaker, circulatory collapse requiring vasopressor or inotropic support, or arrhythmia requiring therapy).
- Cholestatic disorders, or unresolved sinusoidal obstructive syndrome/veno-occlusive disease of the liver (defined as persistent bilirubin abnormalities not attributable to aGvHD and ongoing organ dysfunction).
- Any corticosteroid therapy for indications other than aGvHD at doses >1 mg/kg/day methylprednisolone (or equivalent prednisone dose 1.25 mg/kg/day) within 7 days of Screening. Routine corticosteroids administered during conditioning or cell infusion is allowed.
- Current therapy with medications that interfere with coagulation or platelet function including but not limited to aspirin and related drugs, heparin, and warfarin (to minimize risk of bleeding). Note: Heparin or Low Molecular Weight Heparin was allowed as per protocol amendment 2, if used at sub-therapeutic dose (e.g. for prophylaxis of sinusoidal obstructive syndrome/veno-occlusive disease of the liver).
- History of progressive multifocal leuko-encephalopathy.

- Patients who received JAK inhibitor therapy for any indication after initiation of current alloSCT conditioning.
- Previous participation in a study of any investigational treatment agent within 30 days of randomisation or within 5 half-lives of the investigational treatment agent, whichever is longer.
- Known allergies, hypersensitivity, or intolerance to systemic immunosuppressive therapy.
- Pregnant or nursing (lactating) women.

Grading of GvHD according to MAGIC criteria (Harris et al 2016), is a tool to standardize the collection of complex clinical data for acute GvHD (not advising treatment in clinical practice), and to collect a quantification of symptoms (extent of skin rash, total bilirubin level, volume of diarrhea etc) and still valid and used in most EU transplant centres and recommended by the EBMT.

The choice of aGvHD grading according to Harris et al 2016 is appropriate and presently the most commonly used grading instrument. However, the adherence may vary between centres due to the meticulousness demanded and time-consuming task of grading. The applicant has presented means taken to ensure the reliability and integrity of the data. The high number of PD have been highlighted, however the actual numbers of PD with respect to efficacy assessment for the timepoint of primary and key secondary endpoints were relatively low. Furthermore, a post-hoc sensitivity analysis, based on data from the secondary analysis cut-off, has been conducted excluding patients with PDs related to aGvHD efficacy assessment on the primary and key secondary endpoints showing similar results to those reported in the overall patient population (FAS).

Treatments

Patients were randomized 1:1 to receive either best available therapy (BAT) or ruxolitinib.

Test and reference therapies, dose- and mode of administration:

Patients were randomised to receive:

- RUX 10 mg orally twice daily or
- Best available therapy (BAT) identified by the Investigator prior to patient randomisation among the following treatments: anti-thymocyte globulin (ATG), extracorporeal photopheresis (ECP), mesenchymal stromal cells (MSC), low-dose methotrexate (MTX), mycophenolate mofetil (MMF), mTOR inhibitors (everolimus or sirolimus), etanercept, or infliximab. No other types or combinations of BAT were permitted. Within the first 28 days after randomisation, the initiation or addition of another BAT was allowed for patients meeting criteria of disease progression, mixed response, no response but was considered a treatment failure for both the primary and key secondary objectives.

Dose modifications

RUX dose modifications (reductions and/or interruptions) and delays were permitted to address safety concerns related to haematological (primarily, cytopenias) or non-haematological reasons, or to avoid under- or over-exposure of RUX when combined with strong CYP3A4 inhibitors or dual CYP2C9 and CYP3A4 inhibitors. RUX dose could be re-escalated to a maximum of 10 mg bid when the safety concern was resolved. The permitted daily dose range for RUX was 5 mg to 10 mg bid.

In the BAT arm, dose adjustment was based on manufacturer's instructions, labeling, patient's medical condition, and institutional guidelines.

Treatment duration and tapering

All eligible patients randomised to study treatment were planned to receive treatment from the day of randomisation up to approximately 24 weeks (6 months), or up to approximately 2 years from randomisation, in case the end of RUX taper is delayed due to an aGvHD flare or other safety concerns or the patient met any discontinuation criteria, whichever was earlier.

Tapering of immunosuppression therapy in responding patients was performed in two steps:

1. Taper of corticosteroids: initiated not earlier than Day 7 and performed per institutional guidelines
2. Taper of CNI and/or RUX once the patient stopped corticosteroids as follows: CNI taper was performed as per institutional guidelines. RUX taper was initiated after Day 56, and performed based on condition of the patient, current dosing regimen and clinical judgement of the investigator.

The taper of corticosteroids, CNI, and RUX had to be completed by Week 96.

Special guidelines for cessation or interruption of immunosuppression, in case of an aGvHD flare, were in place.

The variety of therapies in the control arm is acceptable, as a result of differences in national and local guidelines. The most commonly used BAT treatment was ECP. Approximately 20% were treated with more than 1 concomitant BAT treatment.

Initial BAT treatment C2301, REACH 2

Table 14.3-1.3 (Page 1 of 1)
Initial BAT and number of BATs
Safety Set

	BAT N=150
Initial BAT -n (%)	
Extracorporeal photopheresis (ECP)	41 (27.3)
Mycophenolate mofetil (MMF)	25 (16.7)
Etanercept	22 (14.7)
Anti-thymocyte globulin (ATG)	20 (13.3)
Infliximab	17 (11.3)
Mesenchymal stromal cells (MSC)	15 (10.0)
Low-dose Methotrexate (MTX)	5 (3.3)
Sirolimus	3 (2.0)
Everolimus	2 (1.3)
Number of BAT -n (%)	
1	119 (79.3)
2	27 (18.0)
>2	4 (2.7)

CINC424C2301-Secondary CSR

Patients continued to receive systemic corticosteroids with or without continued CNI with or without other systemic treatment for aGvHD per standard of care by the Investigator during the Screening period. Systemic medications for aGvHD other than corticosteroids +/- CNI could be continued after randomization only if used for aGvHD prophylaxis (i.e. started before the diagnosis of aGvHD)

The continued treatment with systemic corticosteroids with or without continued CNI with or without other systemic treatment, after randomisation, is according to standard procedure, and RUX or BAT treatment is considered an add on treatment at randomisation.

Concomitant treatment

Supportive treatments per institutional guidelines for management of alloSCT patients with SR-aGvHD were allowed, including systemic corticosteroids, CNI (cyclosporine or tacrolimus), and topical corticosteroid therapy, however, close monitoring of potential drug-drug interactions effects was recommended. Furthermore, antibiotics, anti-infectives, and immunizations could be used as prophylactic therapies for infections. Dose adjustments of ruxolitinib might be required particularly in patients treated with strong CYP3A4 inhibitors or dual CYP2C9 and CYP3A4 inhibitors due to potential for drug-drug interactions leading to under- or over-exposure.

In line with the MF and PV indications, a 50% dose reduction should be considered in patients with GvHD when co-administered with dual inhibitors of CYP2C9 and CYP3A4 enzymes. The concomitant use of ruxolitinib with fluconazole doses of greater than 200 mg daily was to be avoided. Due to the high risk of bleeding in alloSCT patients with SR-aGvHD, aspirin, NSAIDs, and related medications that would expectedly reduce platelet function and/or heparin, warfarin or related medication that would adversely affect blood coagulation, were prohibited.

A new immunosuppressive agent could be added to RUX or BAT treatment regimen if the patient met the criteria for disease progression, no response, or mixed response, or aGvHD flare failure; this was, then, considered treatment failure.

Objectives

The primary and key secondary objectives were analysed when all patients completed Day 56 or discontinued from study participation earlier with a cut-off date of 25-Jul-2019 and the results were presented in the Primary analysis CSR. The second analysis with a cut-off date of 06-Jan-2020 includes data when all patients have completed approximately 6 months of treatment or discontinued from study participation earlier.

The Primary analysis included the primary endpoint and key secondary endpoint (cut-off date of 25-Jul-2019)

Primary objective was to compare the efficacy of ruxolitinib vs. Best Available Therapy (BAT) in patients with Grade II-IV SR-aGvHD assessed by Overall Response Rate (ORR) at Day 28.

The key secondary objective was to compare the rate of durable ORR (proportion of all patients in each arm who achieved a CR or PR at Day 28 and maintained a CR or PR at Day 56) between RUX and BAT.

Outcomes/endpoints

Summary of efficacy endpoints Study C2301, 271 and D2301

Endpoint	Acute GvHD		Chronic GvHD
	Study C2301 (REACH2)	Study 271 (REACH1)	Study D2301 (REACH3)
Primary endpoint	**Overall response rate (ORR) at Day 28 after randomization**, defined as the proportion of patients in each arm demonstrating a complete response (CR) or partial response (PR) without requirement for additional systemic therapies for an earlier progression, mixed response or non-response. Scoring of response was relative to the organ stage at the time of randomization.	**ORR at Day 28**, defined as the proportion of patients demonstrating a complete response (CR), very good partial response (VGPR), or partial response (PR)	**Overall response rate (ORR) on Cycle 7 Day 1 after randomization**, defined as the proportion of patients in each arm demonstrating a complete response (CR) or partial response (PR) without the requirement of additional systemic therapies for an earlier progression, mixed response or non-response. Scoring of response was relative to the organ score at randomization.
Key secondary endpoint	**Durable ORR at Day 56**, defined as the proportion of all patients in each arm who achieve a complete response (CR) or partial response (PR) at Day 28 and maintain a CR or PR at Day 56.	**Six-month duration of response (DOR)**, defined as the time from first response until GVHD progression or death. DOR was assessed when all patients who are still on study complete the Day 180 visit.	**Failure-Free Survival (FFS)**, composite time to event endpoint incorporating the following FFS events: i) relapse or recurrence of underlying disease or death due to underlying disease, ii) NRM, or iii) addition or initiation of another systemic therapy for cGvHD. **Modified Lee symptom score**, Rate of patients with clinically relevant improvement in modified Lee symptoms score at Cycle 7 Day 1 relative to baseline.
Other secondary endpoints	**ORR at Day 14**, defined as the proportion of patients who achieved ORR (CR+PR) at Day 14.	**ORR at Days 14, 56, and 100**, defined as the proportion of patients demonstrating a CR, VGPR, or PR at Days 14, 56, and 100.	**ORR on Cycle 4 Day 1**, defined as the proportion of patients who achieved overall response (CR+PR) at Cycle 4 Day 1.
	Best Overall Response (BOR), defined as the proportion of patients who achieve Overall response (CR+PR) at any time point up to and including Day 28 and before the start of additional systemic therapy for aGvHD.	**Best ORR**, defined as the proportion of patients demonstrating a CR, VGPR, or PR per the CIBMTR Criteria at any time point and before the start of new anti-aGVHD therapy.	**Best overall response (BOR)**, defined as the proportion of patients who achieved overall response (CR+PR) at any time point (up Cycle 7 day 1 or the start of additional systemic therapy for cGvHD)
	Duration of response (DOR) is assessed for responders only and is defined as the time from first response until aGvHD progression or the date of additional systemic therapies for aGvHD. Onset of cGvHD, or death without prior observation of aGvHD progression are considered as competing risks.	**Three-month DOR**, defined as the time from first response until GVHD progression or death, when all patients who were still on study complete the Day 84 visit.	**Duration of response (DOR)** is assessed for responders only. DOR is defined as the time from first response until cGvHD progression, death, or the date of change/addition of systemic therapies for cGvHD.
	Cumulative steroid dosing until Day 56 Weekly cumulative steroid dose for each patient up to Day 56 or end of treatment was calculated.	**Average and cumulative corticosteroid dose** at Days 28, 56, 100, and 180	Proportion of patients successfully tapered off all corticosteroids at Cycle 7 Day 1
	Non-relapse mortality (NRM), defined as the time from date of randomization to date of death not preceded by hematologic disease relapse/progression.	**Non-relapse mortality (NRM)**, defined as the proportion of patients who died due to cause other than malignancy relapse at months 6, 9, 12, & 24.	**Non-relapse mortality (NRM)**, defined as the time from date of randomization to date of death not preceded by underlying disease relapse/recurrence.
	Overall Survival (OS), defined as the time from the date of randomization to the date of death due to any cause.	**Overall Survival (OS)**, defined as the time from study enrollment (first dose of ruxolitinib treatment) to death due to any cause.	**Overall Survival (OS)**, defined as the time from the date of randomization to the date of death due to any cause.
	Failure-Free Survival (FFS), defined as the time from the date of randomization to date of hematologic disease relapse/progression, NRM or addition of new systemic aGvHD treatment.	**Failure-Free Survival (FFS)**, defined as the time from first dose of ruxolitinib to the earliest date that a patient died, had a relapse/ progression of the underlying malignancy, required additional therapy for aGVHD, or demonstrated signs or symptoms of cGVHD.	FFS is a key secondary and detailed above
	Event-Free Survival (EFS), defined as the time from the date of randomization to the date of hematologic disease relapse/progression, graft failure, or death due to any cause.	NA	NA
	Malignancy Relapse/Progression (MR), defined as the time from date of randomization to hematologic malignancy relapse/progression. Calculated for patients with underlying hematologic malignant disease.	NA	**Malignancy Relapse/ Recurrence (MR)**, defined as the time from date of randomization to hematologic malignancy relapse/recurrence. Calculated for patients with underlying hematologic malignant disease.
	Incidence rate of cGvHD cGvHD, defined as the diagnosis of any cGvHD including mild, moderate, severe.	**Incidence rate of cGVHD** (Days 180 and 365 and overall)	NA

ORR=overall response rate; EOT=end of treatment; OS=overall survival; EFS=event-free survival; FFS=failure free survival; NRM=non-relapse mortality; BOR=best overall response; cGvHD=chronic graft versus host disease; FACT=functional assessment of cancer therapy; BMT=bone marrow transplantation

Efficacy analysis/ Criteria for evaluation

Acute GvHD assessment: Acute GvHD grading was performed by the investigator at every visit during the treatment period and at EOT visit. aGvHD was graded using standard staging criteria for aGvHD (Harris et al 2016): measures of body surface area aGvHD skin rash, stool volumes or frequency per 24 h time period, and serum bilirubin levels, staging by organ (skin; liver; upper gastro-intestinal; lower gastro-intestinal) and overall grading at the time of the evaluation. In addition, biopsy of the organ involved could be performed per institutional practices at Investigator's discretion for aGvHD management. Once randomized, response to study treatment was assessed by the Investigator at every visit during the Treatment Period according to study protocol definition.

CR: defined as a score of 0 for the aGvHD grading in all evaluable organs that indicates complete resolution of all signs and symptoms of aGvHD in all evaluable organs without administration of additional systemic therapies for any earlier progression, mixed response or non-response of aGvHD.

PR: defined as improvement of 1 stage in 1 or more organs involved with aGvHD signs or symptoms without progression in other organs or sites without administration of additional systemic therapies for an earlier progression, mixed response or non-response of aGvHD.

Lack of response was defined as no response, mixed response, or progression.

No response: defined as absence of improvement in any organ involved by aGvHD, without worsening in any involved organ.

Mixed response: defined as improvement of at least 1 stage in the severity of aGvHD in at least one organ accompanied by progression in another organ or development of signs or symptoms of aGvHD in a new organ.

Progression: defined as worsening in 1 or more organs by 1 or more stages without improvement in any involved organ.

Patients were also monitored for occurrence of aGvHD flares occurring during steroid, CNI, and ruxolitinib taper. Additional aGvHD assessment could be performed as per institutional guidelines at investigator's discretion.

Chronic GvHD assessment: Occurrence of definitive and possible manifestations of cGvHD was assessed monthly from Day 1 to Day 56 and at every visit thereafter during the treatment period, at the time of last dose if before Week 24, in responding patients, and at EOT (or Crossover EOT). After EOT (or Crossover EOT), patients were assessed for occurrence of cGvHD at the Safety Follow-up visit if applicable, and at Month 6, at Month 9, at Month 12, at Month 18 and at Month 24 during the Long-Term Follow-up period. Occurrence of cGvHD was not considered an adverse event (AE). cGvHD was graded as per NIH consensus guidelines for cGvHD, as mild, moderate, or severe at the time of cGvHD diagnosis. **Note:** Clinical presentation resembling de novo chronic GvHD or GvHD overlap syndrome with both acute and chronic GvHD features were exclusion criteria, however a mixed picture can occur after randomization.

Graft failure monitoring: Patients were monitored for any evidence of secondary graft failure at each visit from Day 1 during the Treatment, at the time of last dose if before Week 24, in responding patients, at EOT (or Crossover EOT), Safety Follow-up if applicable, and Long- Term Follow-up periods. Occurrence of graft failure was reported as an event and also as an AE. Graft failure was defined as initial whole blood or marrow donor chimerism >5% declining to <5% on subsequent measurements. Donor chimerism was closely monitored to detect graft failure.

Chimerism: Donor chimerism after a hematopoietic stem cell transplant involves identifying the genetic profiles of the recipient and of the donor pre-transplant, and then evaluating the ratio of donor to recipient cells in the recipient's blood, or bone marrow. Chimerism testing using peripheral blood mononuclear cells or bone marrow (or peripheral blood selected CD3+ T cells) was performed during screening (prior to study treatment start), at Day 28 and at Day 56. In addition, for patients who Crossover from BAT to ruxolitinib, chimerism was also performed at Crossover Week 1. Additional chimerism testing could be performed at any time during study (Treatment and Long-Term Follow-up period) at the treating Investigator's discretion according to local institutional practice as indicated.

Hematologic disease relapse/progression assessment: Patients were monitored for any evidence of underlying hematologic disease relapse or progression during the study. Patients were assessed at each visit from Day 1 during the Treatment period, including during crossover period if applicable, at

the time of last dose if before Week 24, in responding patients, at EOT (or Crossover EOT), Safety Follow-Up if applicable, and the Long-Term follow-up period.

- The relapse and progression of the underlying hematologic disease were assessed by the Investigator as per the definitions outlined in protocol. Evaluation and/or evidence of malignancy relapse/progression was conducted according to local institutional practices. Per protocol, study treatment was discontinued underlying hematological disease progression or relapse.

Safety: Safety was monitored by assessing physical examination, vital signs, height and weight, and laboratory assessments including urinalysis, assessment of pregnancy and fertility, clinical chemistry and hematology. Adverse event data was collected at every visit.

Special safety assessments included:

- Pulmonary function test (PFT), if indicated clinically at investigator's discretion per local practices.
- Bleeding, due to the potential complications of thrombocytopenia and/or coagulopathy in the setting of alloSCT.
- Infection monitoring identified as a risk associated with ruxolitinib and BAT for aGvHD therapy.
- Viral reactivation monitoring for hepatitis B and C (HBV viral DNA-PCR and HCV RNA-PCR), Cytomegalovirus (CMV viral DNA quantification), Epstein Barr Virus (EBV viral load), Human Herpes virus (HHV-6 viral load).
- Second primary malignancy monitoring defined as any new malignancy other than the underlying hematologic disease.

Pharmacokinetics: Blood sampling for PK of ruxolitinib was performed in all patients enrolled in the study and treated with ruxolitinib to characterize the PK parameters in aGvHD patients.

Extensive PK sampling schedule: "Extensive PK" sampling schedule was followed for approximately the first twenty-five (25) adult patients and all adolescent patients enrolled. The 'Extensive PK' sampling scheme includes a pre-dose and seven (7) post-dose samples on Day 1 and Day 7 thereafter, two (2) samples (1 pre-dose and 1 post-dose) per scheduled visit.

Sparse PK sampling schedule: Adult patients randomized to ruxolitinib after the Extensive PK samples were collected, and any patients crossing over from BAT to ruxolitinib after Day 28, would follow the "Sparse PK" sampling schedule and had a total of two (2) samples (1 pre-dose and 1 post-dose) per scheduled visit. The plasma samples from all patients were assayed for ruxolitinib concentrations using validated liquid chromatography-tandem mass spectrometry method.

Resource Utilization was captured for use in post-study Health Economics analysis.

Patient-reported outcomes: In order to measure Quality-of-Life (QoL) among aGVHD patients, and potential changes over time, two patient-reported outcome (PRO) instruments were administered: FACT-BMT and EQ-5D-5L. FACT-BMT is a 50-item self-report questionnaire that measures in adult patients only, the effect of a therapy on domains including physical, functional, social/family, and emotional wellbeing, together with additional concerns relevant for bone marrow transplantation patients. EQ-5D-5L descriptive classification consists of five dimensions of health: mobility, self-care, usual activities, anxiety/depression, and pain/discomfort (Brooks 1996). Each dimension of health has five levels of severity (Herdman et al 2011).

These PRO instruments were planned to be administered on randomization day and every week during the first 2 months, and every 4 weeks thereafter until the end of treatment (EOT).

PRO instruments were not used for adolescents.

Sample size

A study with a total of 308 patients and 1:1 randomization (ruxolitinib vs. BAT) stratified on aGvHD grade (Grade II vs. Grade III vs. Grade IV) would have 90% power to test for the primary endpoint (ORR at Day 28) and approximately 90% power to test for the key secondary endpoint (durable ORR at Day 56). The family wise α-level would be controlled at 0.025 overall for the two comparisons. Specifically, this study would claim to have achieved the efficacy objective when the primary endpoint ORR at Day 28 shows a significant treatment effect at one-sided α = 0.025. Conditional on significance of the primary endpoint, the key secondary endpoint durable ORR at Day 56 would be tested at one-sided α = 0.025.

The sample size calculation was based on the primary variable ORR at Day 28. Based on Martin P. et al BBMT 2013, the ORR at Day 28 in the BAT arm was expected to be 58%. The stratum specific rates (Grade II 69%, Grade III 59%, Grade IV 50%) were obtained assuming that the ratio of aGvHD Grade II: III: IV is 0.2:0.4:0.4. It was expected that treatment with ruxolitinib would result in an 18% increase in the ORR, i.e., an expected odds ratio of 2.25 (which corresponds to an increase in ORR to 75%). Power for the CMH test, stratifying on aGvHD grade, was calculated using software package East V6. In order to ensure 90% power a total sample size of 308 patients is needed. With a sample size of 154 patients in each treatment arm, an observed odds ratio greater than or equal to 1.63 would achieve statistical significance for the primary endpoint. Assuming that the observed response rates in Grades II/III/IV in BAT arm are 69%/59%/50% (overall 57%), observed response rates ≥ 78%/70%/62% (overall 68%) in the ruxolitinib arm would achieve statistical significance.

Based on [Van Groningen 2016], the durable ORR at Day 56 in the BAT arm was expected to be approximately 35%. It is expected that treatment with ruxolitinib would result in a 20% increase in the ORR, i.e., an expected odds ratio of 2.25 (which corresponds to an increase in durable ORR to 55%). With these assumptions and sample size of 308 patients, the power for the key secondary endpoint is at least 90%. With sample size 154 patients in each treatment arm, an observed odds ratio greater than or equal to 1.59 for durable ORR at Day 56 would achieve statistical significance. Assuming that the observed durable response rates in Grades II/III/IV in BAT arm are 45%/36%/30% (overall 35%), observed durable response rates ≥ 57%/47%/41% (overall 47%) in the ruxolitinib arm would achieve statistical significance.

Randomisation

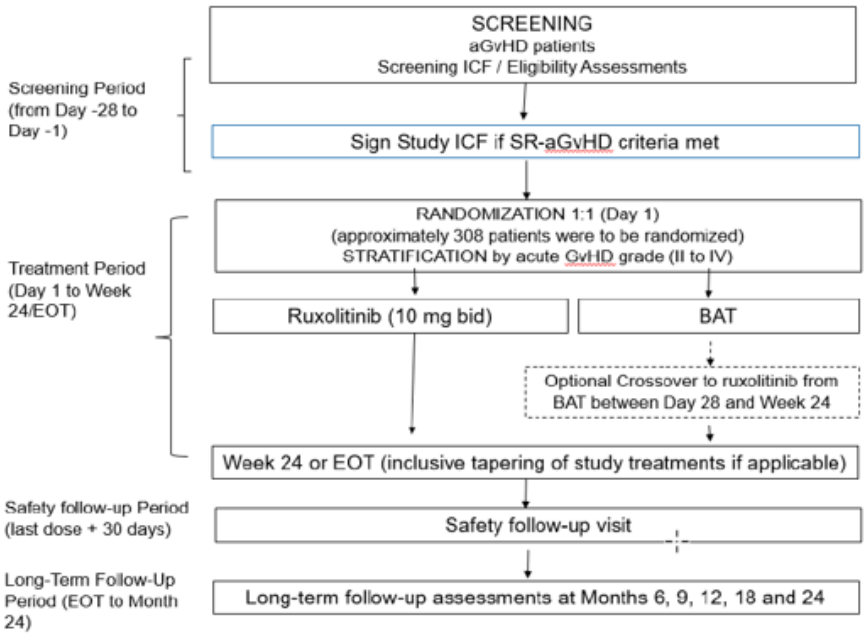
A stratified 1:1 randomisation was conducted, and patients received either ruxolitinib or BAT. They were stratified by aGvHD grade (Grade II vs. III vs. IV). Before randomization, BAT was selected by the Investigator.

Blinding (masking)

As these BATs vary from administered tablets to cellular therapy and photopheresis, the open label design of these studies was inevitable. This has been acknowledged by the CHMP at the time of the Scientific Advice sought in 2016 for cGvHD (CHMP Scientific Advice). The open label strategy is, therefore, acceptable.

The BAT treatment was allocated before randomization took place, i.e., the treatment/s which would be standard treatment/s at the specific site.

Treatment phases/study conduct



Screening period

Definition of steroid refractory aGvHD:

During the screening period, patients were monitored for a diagnosis of Steroid-refractory aGvHD, which was defined as patients who had high-dose systemic corticosteroids (methylprednisolone 2 mg/kg/day [or equivalent prednisone dose 2.5 mg/kg/day]), given alone or combined with CNI, who either:

a) Progressed based on organ assessment after at least 3 days compared to organ stage at the time of initiation of high-dose systemic corticosteroid +/- CNI for the treatment of Grade II-IV aGvHD,

OR

b) Failed to achieve at a minimum a partial response based on organ assessment after 7 days compared to organ stage at the time of initiation of high-dose systemic corticosteroid +/- CNI for the treatment of Grade II-IV aGvHD,

OR

c) Failed corticosteroid taper defined as fulfilling either one of the following criteria:

- Requirement for an increase in the corticosteroid dose to methylprednisolone ≥ 2 mg/kg/day (or equivalent prednisone dose ≥ 2.5 mg/kg/day)

OR

- Failure to taper the methylprednisolone dose to <0.5 mg/kg/day (or equivalent prednisone dose <0.6 mg/kg/day) for a minimum 7 days.

The Screening Phase should not exceed 28 days. The period from suspicion of SR-aGvHD and urgent need of additional treatment is usually short and treatment delay is not desirable.

Randomized treatment phase

Duration of treatment: The End of Treatment (EOT) visit occurred when the patient completed the study treatment period or earlier if the patient met any of the criteria for discontinuation of study treatment.

Patients' treatment period was up to 6 months (Week 24). However, ruxolitinib taper could be delayed up to 2 years from randomization due to an aGvHD flare or other safety concerns.

Patients meeting the primary endpoint at Day 28 will continue ruxolitinib at 10 mg BID until Day 56. Ruxolitinib will then be tapered in responding patients who have completed steroid taper starting Day 56 to complete treatment by Week 24, while patients responding to BAT will be managed as per institutional practices.

For patients not meeting the primary endpoint at Day 28: Patients who are randomized to ruxolitinib, will discontinue treatment and be treated per Investigator's judgement.

Crossover treatment phase

Patients who are randomized to BAT, and who meet cross-over criteria after Day 28, may cross over to the ruxolitinib treatment arm and follow the same treatment duration and taper schedule as patients originally randomized to ruxolitinib treatment.

Corticosteroids and CNI for aGvHD treatment are allowed to be continued, with cessation required of any other systemic immunosuppressive treatment prior to crossover.

Cross-over to ruxolitinib during the Treatment Period for patients randomized to BAT was an option between Day 28 and Week 24 if they:

- Failed to meet the primary endpoint response definition (CR or PR) at Day 28

OR

- Lost the response thereafter and met criteria for progression, mixed response, or no response, necessitating new additional systemic immunosuppressive treatment for aGvHD.

AND

- Did not have signs/symptoms of chronic Graft vs. Host Disease (cGvHD) (overlap syndrome, progressive, or de novo cGvHD). Patients who crossed over to ruxolitinib were followed until completion of treatment with ruxolitinib and received the same treatment and tapering schedule as patients randomized to ruxolitinib treatment.

Safety Follow-Up (last dose + 30 days)

A 30-day Safety Follow-up visit was performed for all patients after the last dose of ruxolitinib (administered during Treatment Period or after Crossover) or BAT.

Long-Term Follow-Up Period (From EOT to Month 24)

As SR-aGvHD often leads to death within 2 years after the transplant, and assessment of long-term safety and durable efficacy is clinically relevant, all patients (responders and non-responders in both arms, regardless of when treatment was discontinued) would be followed up for long-term observation up to 24 months from randomization.

During this period, long-term data was collected: survival, any relapse/progression of the underlying hematologic disease for which the alloSCT procedure was performed, non-relapse mortality (NRM), any occurrence of graft failure, event-free survival (EFS), any occurrence of cGvHD, and occurrence of any second primary malignancies. Visits for these assessments occurred after EOT or Crossover EOT, at 6, 9, 12, 18, and 24 months.

Statistical methods

Multiple analyses were performed:

- The first analysis of the primary and key secondary endpoints was performed with the cut-off date as 25-Jul-2019 after all patients have completed their Day 56 visit or have discontinued study. The results were summarized in the so called "primary" clinical study report (CSR).
- Further analyses on secondary endpoints were performed when all patients have completed 6 months treatment or discontinued from study participation earlier, using all data collected in the database up to the data cut-off at 06-Jan-2020.

The final analysis will occur once all patients have completed the study (up to 24 months from randomization). All available data from all patients up to end of study will be reported in a final CSR. No formal interim efficacy analysis was planned in this study.

Analysis sets

Full Analysis Set (FAS) comprises all patients to whom study treatment has been assigned by randomization. According to the intent to treat principle, patients were analysed according to the treatment and strata they have been assigned to during the randomization procedure.

The Per-Protocol Set (PPS) consists of a subset of the patients in the FAS who are compliant with requirements of the clinical study protocol. The following list of protocol deviations lead to exclusion of the patient from the PPS:

- Not steroid refractory aGvHD
- More than one prior systemic therapy for the treatment of aGvHD other than corticosteroids +/- CNI (prophylaxis or treatment)
- Missing or incorrect aGvHD grade at randomization
- Taking any prohibited medication as specified in the protocol after start of study treatment and before end of study treatment
- Study treatment received different from treatment arm assigned by randomization.

The Safety Set includes all patients who received at least one dose of study treatment. Patients were analysed according to the study treatment received, where treatment received was defined as the randomized treatment if the patient took at least one dose of that treatment, or the first treatment received if the randomized treatment was never received during the randomized treatment period.

The Crossover Analysis Set (CAS) comprises all patients randomized to and received BAT who then crossed over and received at least one dose of ruxolitinib. This analysis set was used for all analyses for crossover patients.

Primary analysis

The primary endpoint, ORR at Day 28, was defined as the proportion of patients with CR or PR at Day 28 according to standard criteria [Harris 2016]. Note that response is relative to the assessment of aGvHD at randomization. Lack of response is defined as no response, mixed response, or progression.

For definitions of these events, see section Outcomes/Endpoints. A patient was not considered a responder at Day 28 if any of the following events occurs:
- Missing aGvHD assessment at baseline or Day 28
- No CR or PR at Day 28
- Additional systemic therapy for aGvHD prior to Day 28.

The primary analysis was performed at the time when all patients have completed their Day 56 visit or discontinued earlier. The Cochrane-Mantel-Haenszel (CMH) chi-square test, stratified by the randomization stratification factor (i.e., aGvHD Grade II vs III vs. IV), was used to compare ORR between the two treatment groups, at the one-sided 2.5% level of significance, using the analysis the FAS. The statistical hypotheses tested were:

H0: ORRrux ≤ ORRBAT vs. H1: $ORR_{rux} > ORR_{BAT}$

where ORR_{rux} and ORR_{BAT} are the overall response rates at Day 28 in the ruxolitinib and BAT group, respectively.

One-sided p-value, odds ratio and 95% Wald confidence limits calculated from the stratified CMH test were presented. ORR was also summarized using descriptive statistics (N, %) by treatment arm along with two-sided exact binomial 95% CIs [Clopper and Pearson 1934].

As a supportive analysis, the primary endpoint was analysed in the same way as in the primary analysis based on patients in PPS. A supportive analysis was also conducted using logistic regression model to estimate the treatment effect adjusting for key baseline and prognostic factors based on patients in FAS. The model included the followings covariates: age, gender, race, aGvHD grade, source of grafts, criteria for SR-aGvHD, prior aGvHD therapy in addition to treatment as one of the covariates.

As a sensitivity analysis, the two treatment groups were compared using Fisher's exact test.

Subgroup analyses were performed for the primary efficacy endpoint to examine the homogeneity of treatment effect. For each of the subgroups, the following statistics were presented:
- Proportion of patients with ORR using descriptive statistics (N, %) along with two-sided exact binomial 95% CIs.
- Odds ratio with 95% CI from a logistic regression model with treatment and stratification factors as covariates.

Efficacy analyses in subgroups were purely exploratory and were intended to explore the consistency of treatment effect. No inferential statistics (p-values) were to be produced for the subgroups.

ORR at Crossover Day 28

ORR at Crossover Day 28 was defined as proportion of crossover patients with CR or PR at Crossover Day 28. Note that response is relative to the last assessment of aGvHD prior to or at the start date of crossover treatment (ruxolitinib). A patient was not considered a responder at Crossover Day 28 if any of the following events occurs:
- Missing aGvHD assessment at Crossover baseline or Crossover Day 28
- No CR or PR at Crossover Day 28
- Additional systemic therapy for aGvHD prior to Crossover Day 28.

ORR at Crossover Day 28 was summarized using descriptive statistics (N, %) along with two-sided exact binomial 95% CIs [Clopper and Pearson 1934] based on the CAS.

Analysis of the key secondary endpoint

The key secondary endpoint was Durable ORR at Day 56, an tested only if the primary endpoint was statistically significant. If a patient was a CR at Day 28 and a PR at Day 56, he/she was considered as

a durable responder. A patient was not considered a durable responder at Day 56 if any of the following events occurs:
- Not a responder at Day 28
- Missing aGvHD assessment at Day 56
- No CR or PR at Day 56.
- Additional systemic therapy for aGvHD prior to Day 56.

The patients randomized to BAT who met cross-over criteria and crossed-over to ruxolitinib were considered to have the additional systemic therapy for aGvHD and were not considered as a responder afterwards. Durable ORR at Day 56 was analyzed using the same method as in the primary analysis.

Durable ORR at Crossover Day 56

Durable ORR at Crossover Day 56 was defined as the proportion of all crossover patients who achieved a CR or PR at Crossover Day 28 and maintained a CR or PR at Crossover Day 56. Note that response is relative to the last assessment of aGvHD prior to or at the start date of crossover treatment (ruxolitinib). A patient was not considered a durable responder at Crossover Day 56 if any of the following events occurs:
- Not a responder at Crossover Day 28
- Missing aGvHD assessment at Crossover Day 56
- No CR or PR at Crossover Day 56.
- Additional systemic therapy for aGvHD prior to Crossover Day 56.

Durable ORR at Crossover Day 56 was summarized using descriptive statistics (N, %) along with two-sided exact binomial 95% CIs [Clopper and Pearson 1934] based on the CAS.

Analyses of other secondary endpoints

Other secondary endpoints were not part of the hierarchical testing. ORR at Day 14 and BOR were analysed only at the data cut-off date of 25-Jul-2019. The rest of the endpoints were re-analysed at the second analysis, i.e. when all patients have completed 6 months treatment or discontinued from study participation earlier (data cut-off date 06-Jan-2020). These secondary efficacy endpoint analyses were non-comparative in nature and analysed based on the FAS.

Safety and PRO endpoints were analysed using descriptive statistics.

Handling of missing values/censoring/discontinuations

Patients with missing assessments that prevent the evaluation of the primary endpoint were considered non-responders on that treatment arm. This includes missing aGvHD response assessments at baseline and Days 28 and 56. Duration of response was not censored based on a treatment discontinuation. No data imputation was to be applied.

Multiplicity

The primary and the key secondary endpoint were tested hierarchically. That is, if the ORR at Day 28 was statistically significant, the durable ORR at Day 56 would be tested.

The two endpoints were tested hierarchically controlling alpha at 2.5 % one-sided. As this is the only analysis of these two endpoints and included all patients, no further adjustment for multiplicity was required then the hierarchical procedure used.

Data Monitoring Committee (DMC)

A DMC has been added as a part of study protocol amendment 2, in order to limit potential bias in study conduct in the context of this ongoing open label trial. The DMC is responsible for assessing the safety data defined as SAEs/death and graft failures obtained in the study and ensuring that the event rates are balanced between the treatment arms.

Changes from the planned analyses

Compared to the protocol version 00 (dated 01-Sep-2016), 'discontinuation from randomized treatment' has been dropped as reason. Since some BATs have fixed duration which is less than 8 weeks, to avoid bias, discontinuation from randomized treatment will not be used to disqualify the response or durable response. Also, duration of response was not censored based on a treatment discontinuation.

According to the protocol version 00, patients who discontinued randomized treatment prior to or at Day 28 would be considered as ORR non-responders. Similar rule was to be applied for the durable ORR for the period prior to or at Day 56.

There were very few substantial changes from the protocol specified analysis, and the analyses were actually performed in accordance to the original protocol (version 00). Handling of discontinuations from randomized treatment, which was identified as a change from the protocol specified analysis, is addressed in a separate question. It was noted that at the time of the second analysis (data cut-off date 06-Jan-2020), due to the data cleaning activities, one subject in each treatment arm has changed from a non-responder to responder at Day 28. Although there is no change to the odds ratio and the conclusion, it was of concern at what time-period the data cleaning was performed for the first (and second) analysis. Dates for the database locks, data cleaning period and SAP versions are provided for the two data cut-offs.

According to the SAP, the one change to the analysis described in the original protocol was that treatment discontinuations would not be used to disqualify the response or durable response, which is not obvious to have been followed as all treatment discontinuations were treated as non-responders in the primary and the key secondary analysis. Among non-responders, the rates of early discontinuations appear to be well-balanced between the treatment groups, but together with death and missing visits, this may add to the uncertainty in the estimated outcome.

Of note, for the interpretation of the phase III study results, the convention is to use 2-sided tests and p-values. For that reason, 2-sided p-values are included in the SmPC.

Protocol amendments

There were 2 protocol amendments. The key features of each amendment are presented below:

Amendment 1 (31- May-2017): The main purpose of Amend 1 was to clarify the exclusion criterion #5 and other eligibility criteria to follow standard medical practice. At the time of randomization, the investigator assessed if the patient presented with a viral infection or not based on his/her medical judgement and without waiting for viral load test results for CMV, EBV, HHV-6, HBV, HCV performed at screening in order not to delay the initiation of aGvHD treatment in this life-threatening condition.

Amendment 2 (21- Jun-2018): The main purpose of the amendment was to allow for more flexibility in the tapering of corticosteroids, calcineurin inhibitors (CNI) and ruxolitinib; and if needed, for this taper to be completed safely beyond Week 24. This change included clarification that institutional guidelines for the tapering of corticosteroids and CNI could be followed. Additionally, the physician could tailor the tapering strategy to each patient's condition, including stopping ruxolitinib more slowly in case of an acute Graft vs. Host Disease (aGvHD) flare or other safety concerns which may prevent the taper from

being completed by Week 24. Patients, who met the protocol criteria for treatment discontinuation were not eligible to continue receiving ruxolitinib within the study but were given the possibility to continue/receive ruxolitinib outside the study (provided they met specified criteria), would not enter the Long-Term Follow-Up period. -Alignment with medical practices in managing adolescent patients, other systemic medications for aGvHD prophylaxis in addition to CNI could be maintained after randomization for all patients. To be allowed, these additional prophylactic medications start date must precede the diagnosis date of aGvHD. This change is anticipated to primarily improve adolescent enrolment but its impact on overall patient homogeneity is limited. -Simplifying the PK sampling. -Alignment of the secondary endpoint BOR to aGVDH publications. A DMC was added to address the Study Steering Committee's request to be informed on the balance of safety events between treatment arms. The DMC was added to maintain the Study Steering Committee blinding during their review of pooled safety data, in order to preserve the integrity of the trial.

CSR amendment

In addition to protocol amendments above, the Clinical Study Report was amended 12-Aug-2020 due to the following reasons

- Correction of the cumulative corticosteroids dosing results in the primary analysis CSR because the doses of methylprednisolone were not converted to equivalent prednisone doses. The conversion factor (1.25 to convert methylprednisolone dose in prednisone) was mistakenly omitted in the programming data specification document (PDS) and statistical analysis plan (SAP), which now has been added. Correction and update of the biochemistry results
- Furthermore, Correction and update of the lipase biochemistry results has been added and a conversion issue was detected for creatinine, magnesium and phosphate where few cases were reported as CTCAE grade 1 instead of grade 0. Some laboratory listings have been reformatted for more clarity and a clarification was added as a footnote in the outputs to a category of a protocol deviation related to the screening informed consent in order to align with its definition in the study specification documents.

The protocol amendments 1 and 2 are not expected to interfere with the integrity of the study. It is not expected that the amendments of the CSR, with regard to the recalculation of corticosteroids and biochemistry data, will affect the interpretation of data.

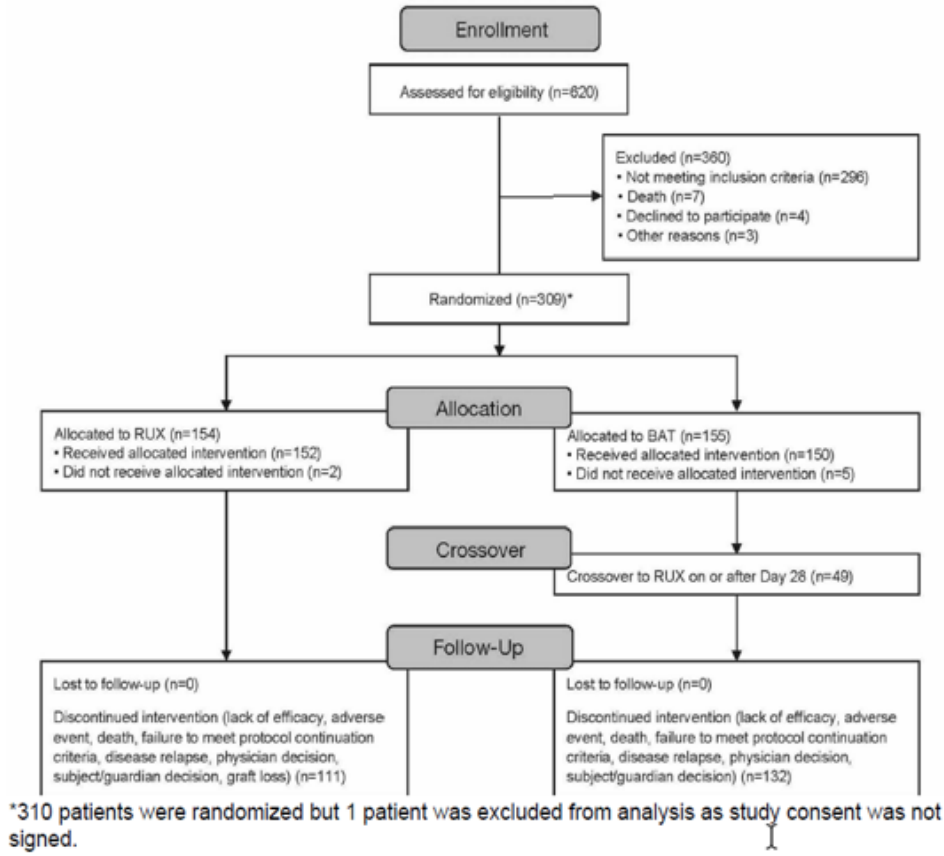
Results

Recruitment

Table 14.1-2.1 (Page 1 of 1)
Screening phase disposition
All Screened Subjects

	All Subjects N=620 n (%)
Screened	620 (100)
Completed screening phase, randomized	309 (49.8)
Completed screening, not randomized	1 (0.2)
Did not complete screening	310 (50.0)
Primary reason for not completing screening	
Screen failure	296 (47.7)
Death	7 (1.1)
Subject/guardian decision	4 (0.6)
Physician decision	2 (0.3)
Technical problems	1 (0.2)
Adverse event	0
Disease relapse	0
Failure to meet protocol continuation criteria	0
Graft loss	0
Lack of efficacy	0
Lost to follow-up	0
Pregnancy	0
Protocol deviation	0
Study terminated by sponsor	0

Participant flow



A high rate of screen failures was seen, 310 out of a total of 620 aGvHD screened patients (50%) did not complete screening. A majority of these were screen failures, with the primary reason for not

meeting eligibility criteria was that patients did not become steroid refractory (42.9%) according to study criteria.

Due to the low acceptance of a lag period in these patients from the time of SR-aGvHD has occurred until treatment should start, screening activities and assessment of inclusion and exclusion criteria began once the patient was diagnosed with aGvHD, i.e., before a possible conversion to SR-aGvHD. Out of the 309 patients included in the FAS, 154 were in the ruxolitinib arm (including 5 adolescent patients) and 155 in the BAT arm (including 4 adolescent patients).

Any occurrence of SR-aGvHD was monitored closely.

Disposition of patients

Study treatment was discontinued due to following reasons: Lack of efficacy of aGvHD treatment, i.e. not achieving PR or CR at Day 28 or Crossover Day 28 and/or requiring additional systemic therapy for aGvHD, at any time; development of signs or symptoms of cGvHD including de novo, overlap, or progressive onset; underlying haematological disease progression or relapse; evidence of graft failure necessitating rapid taper of immunosuppression, administration of non-scheduled DLI, stem cell boost, chemotherapy, or other treatment that would expectedly affect aGvHD; adverse events leading to study treatment discontinuation.
A majority of patients discontinued from the randomized treatment phase (24 weeks) in both arms. The median duration of exposure to RUX treatment and BAT treatment were 63 days (6.0 – 463.0 days) and 29 days (1.0 - 188.0 days) respectively.

Randomized treatment phase at the secondary analysis, data cut-off 06 Jan 2020

Table 10-1 Patient disposition (FAS)

	RUX N=154 n (%)	BAT N=155 n (%)	All Patients N=309 n (%)
Subjects randomized			
Treated	152 (98.7)	150 (96.8)	302 (97.7)
Not treated	2 (1.3)	5 (3.2)	7 (2.3)
Treatment ongoing*	3 (1.9)	0	3 (1.0)
Completed treatment period	35 (22.7)	21 (13.5)	56 (18.1)
Discontinued from treatment period	116 (75.3)	134 (86.5)	250 (80.9)
Reason for discontinuation			
Lack of efficacy	32 (20.8)	69 (44.5)	101 (32.7)
Adverse event	27 (17.5)	6 (3.9)	33 (10.7)
Death	25 (16.2)	21 (13.5)	46 (14.9)
Failure to meet protocol continuation criteria	12 (7.8)	9 (5.8)	21 (6.8)
Disease relapse	8 (5.2)	13 (8.4)	21 (6.8)
Physician decision	6 (3.9)	9 (5.8)	15 (4.9)
Subject/guardian decision	4 (2.6)	6 (3.9)	10 (3.2)
Graft loss	2 (1.3)	0	2 (0.6)
Lost to follow-up	0	0	0
Pregnancy	0	0	0
Protocol deviation	0	0	0
Study terminated by sponsor	0	0	0
Technical problems	0	1 (0.6)	1 (0.3)
Continued into next phase at the end of randomized treatment			
Crossover treatment	0	49 (31.6)	49 (15.9)
Entered long-term follow-up	96 (62.3)	51 (32.9)	147 (47.6)

* Ongoing at the time of the data cut-off date 06-Jan-2020

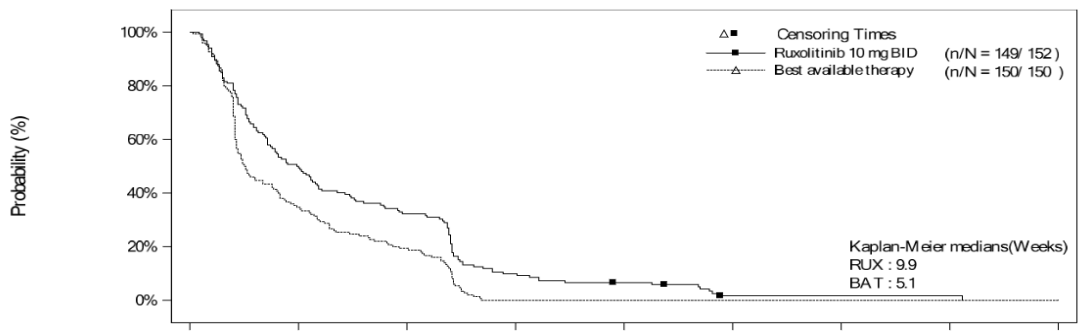
The study had three phases: Randomization treatment phase (n=309), crossover treatment phase (n=49) and long-term follow-up phase into which 87 patients entered (56.5%) in the RUX arm and 45 (29.0%) patients in the BAT arm. In addition, 24 patients from the cross-over phase entered the long-term follow-up phase.

In the randomization treatment phase, thirty-five (22.7%) patients completed the treatment period in the RUX arm compared to 21 (13.5%) in the BAT arm. A total of 116 (75.3%) patients in RUX arm and 134 (86.5%) patients in BAT arm discontinued during the randomized treatment period and the most common reasons for discontinuation were lack of efficacy (20.8% in ruxolitinib arm vs. 44.5% in BAT arm), death (16.2% in ruxolitinib arm vs. 13.5% in BAT arm) and AEs (17.5% in ruxolitinib arm vs. 3.9% in BAT arm). Among the 14 patients (6 in ruxolitinib arm and 8 in BAT arm) who discontinued due to physician decision, one patient in ruxolitinib arm and two patients in BAT arm were reported to be responders. In the RUX arm, 62.3% entered the long-term follow up phase compared to 32.9% in the BAT arm.

Due to the option for patients in the BAT arm to switch to a second BAT within the first 28 days (i.e., the timepoint for the analysis of the primary endpoint and also the option for BAT patients to crossover to RUX treatment), without discontinuing from the main treatment period, indicate no difference between treatment arms until this timepoint but a quick drop, for the BAT arm, was thereafter seen.

The K-M plot over treatment discontinuations indicate similar and rather quick drop out in both treatment arms already from the study start, and a sharp drop in the BAT arm after day 28.

Figure HA2 14.1-1.5 (Page 1 of 1)
Kaplan-Meier estimate of time to discontinuation from main study period
Safety set



No. of patients still at risk

Time(Weeks)	0	10	20	30	40	50	60	70	80
Ruxolitinib 10 mg BID	152	76	49	15	9	1	1	1	0
Best available therapy	150	52	29	0	0	0	0	0	0

The most frequent reasons for discontinuations during the first 28 days were lack of efficacy, death, and adverse events in ruxolitinib arm (13, 12 and 7 patients, respectively), and lack of efficacy and death in the BAT arm (26 and 9 patients, respectively). The sharp drop after day 28 was not surprising since per study design patients were expected to remain in the initial treatment at least until Day 28.

CINC424C2301-Secondary CSR

Table HA2 14.1-1.4
Discontinuation reasons by time interval for the main study period, by treatment
Full analysis set

Treatment: Ruxolitinib 10 mg BID

	Days since treatment start								
Reason	**<=28**	**>28 - 56**	**>56 - 112**	**>112 - 168**	**>168 - 336**	**>336 - 672**	**>672**	**Not treated**	**any**
Primary reason for discontinuation									
Adverse event	7	9	7	3	1	0	0	0	27 (17.5%)
Death	12	7	5	1	0	0	0	0	25 (16.2%)
Disease relapse	1	1	3	2	1	0	0	0	8 (5.2%)
Failure to meet protocol continuation criteria	0	2	3	3	4	0	0	0	12 (7.8%)
Graft loss	0	1	1	0	0	0	0	0	2 (1.3%)
Lack of efficacy	13	12	6	1	0	0	0	0	32 (20.8%)
Physician decision	0	3	3	0	0	0	0	0	6 (3.9%)
Subject/guardian decision	0	1	0	0	1	0	0	2	4 (2.6%)
Technical problems	0	0	0	0	0	0	0	0	0

CINC424C2301-Secondary CSR

Table HA2 14.1-1.4
Discontinuation reasons by time interval for the main study period, by treatment
Full analysis set

Treatment: Best available therapy

	Days since treatment start								
Reason	**<=28**	**>28 - 56**	**>56 - 112**	**>112 - 168**	**>168 - 336**	**>336 - 672**	**>672**	**Not treated**	**any**
Primary reason for discontinuation									
Adverse event	3	0	2	0	1	0	0	0	6 (3.9%)
Death	9	5	5	0	0	0	0	2	21 (13.5%)
Disease relapse	2	2	5	4	0	0	0	0	13 (8.4%)
Failure to meet protocol continuation criteria	0	1	3	3	1	0	0	1	9 (5.8%)
Graft loss	0	0	0	0	0	0	0	0	0
Lack of efficacy	26	30	8	5	0	0	0	0	69 (44.5%)
Physician decision	3	3	2	0	0	0	0	1	9 (5.8%)
Subject/guardian decision	3	1	0	1	0	0	0	1	6 (3.9%)
Technical problems	1	0	0	0	0	0	0	0	1 (0.6%)

No patients discontinued due to protocol deviations.

Cross over patients

Forty-nine patients (31.6%) discontinued BAT and crossed over to RUX treatment of whom 55.1% entered the long-term follow-up phase. The median time to cross-over was 34 days (range 28.0-162.0). At the secondary analysis data cut-off, 36 (73.5%) of patients discontinued the cross-over treatment period. The most common reason for discontinuation were AEs (24.5%), death (16.3%) and lack of efficacy (12.2%).

Conduct of the study

Protocol deviations

Table 14.1-4.1 (Page 1 of 2)
Protocol deviations
Full Analysis Set

Category Protocol deviation	RUX N=154 n (%)	BAT N=155 n (%)	All Patients N=309 n (%)
Any protocol deviation	131 (85.1)	135 (87.1)	266 (86.1)
Any exclusion criteria deviation	4 (2.6)	3 (1.9)	7 (2.3)
exclusion criteria of concomitant meds not met	2 (1.3)	0	2 (0.6)
Exclusion criterion of absence of chronic GvHD(denovo or overlap syndrome) met.	1 (0.6)	1 (0.6)	2 (0.6)
exclusion criteria of absence of significant or uncontrolled cardiac disease not met	1 (0.6)	0	1 (0.3)
Exclusion criteria of absence of severely impaired renal function not met	0	1 (0.6)	1 (0.3)
exclusion criteria of absence of relapsed primary maligancy not met	0	1 (0.6)	1 (0.3)
Any inclusion criteria deviation	56 (36.4)	61 (39.4)	117 (37.9)
Viral Load assessment at screening or Day 1 is beyond protocol acceptable window.	26 (16.9)	25 (16.1)	51 (16.5)
Chimerism at Screening or Day 1 is beyond protocol acceptable window.	20 (13.0)	18 (11.6)	38 (12.3)
Study Informed Consent not obtained	8 (5.2)	7 (4.5)	15 (4.9)
Steroid refractory aGvHD criteria not met	7 (4.5)	15 (9.7)	22 (7.1)
acute GvHD grade criteria not met	6 (3.9)	3 (1.9)	9 (2.9)
Myeloid platelet engraftment not confirmed	3 (1.9)	6 (3.9)	9 (2.9)
Screening Informed Consent not obtained	2 (1.3)	0	2 (0.6)
Other Deviation	95 (61.7)	101 (65.2)	196 (63.4)
OTHER DEVIATION	95 (61.7)	101 (65.2)	196 (63.4)
Prohibited Concomitant Medication	31 (20.1)	28 (18.1)	59 (19.1)

Numbers (n) represent counts of subjects.
A Subject can have more than one deviation.

/csr_1/pgm/saf/tprtdev.sas@@/main/4-23JAN2020:16:44 Final

Table 14.1-4.1 (Page 2 of 2)
Protocol deviations
Full Analysis Set

Category Protocol deviation	RUX N=154 n (%)	BAT N=155 n (%)	All Patients N=309 n (%)
Prohibited Concomitant Medication			
prohibited concomitant medication	31 (20.1)	28 (18.1)	59 (19.1)
Subject Not Withdrawn As Per Protocol	10 (6.5)	19 (12.3)	29 (9.4)
Study treatment continued whereas a withdrawal criterion was met	6 (3.9)	0	6 (1.9)
study treatment withdrawal criteria not met	5 (3.2)	19 (12.3)	24 (7.8)
Treatment Deviation	44 (28.6)	60 (38.7)	104 (33.7)
Investigational study treatment dispensing error	43 (27.9)	60 (38.7)	103 (33.3)
criteria for dose reduction / interruption not followed	3 (1.9)	3 (1.9)	6 (1.9)

The high rate of protocol deviations did, overall, not differ between the 2 study arms. The most common deviation belonged to the category "Other deviations" (61.7% and 65.2%) which included: Different aGvHD overall grades used for randomization between e-CRF and IRT; missing two consecutive monthly scheduled viral load testing; missing aGvHD assessment at day 28 or day 56; organ Staging assessment done per investigator criteria/judgement rather than MAGIC criteria (Harris et al); response assessment done per investigator criteria/judgement rather than protocol definition; protocol amendment 2 implemented before patient's re-consent obtained. "Prohibited drugs" were taken by 19.1% patients (20.1% in ruxolitinib arm and 18.1% in BAT arm). Deviations in these categories were the most common reasons for exclusion of patients from the PPS (see below).

With the exception of protocol deviation 'organ staging assessment done per investigator criteria/judgement rather Harris' (18.8% with RUX vs. 34.2% with BAT), no imbalances in

subcategories of 'other deviations' have been observed. It is acknowledged that above-described exception was reported with similar rates at key time points (baseline, Day 28 and Day 56).

Baseline data

Demographics and baseline characteristics (FAS)

Demographic variable	RUX N=154	BAT N=155	All Patients N=309
Age (years)			
n	154	155	309
Mean (SD)	48.1 (16.30)	50.9 (14.97)	49.5 (15.69)
Median	52.5	54.0	54.0
Q1-Q3	32.0-61.0	41.0-63.0	38.0-62.0
Min - Max	12.0-73.0	13.0-71.0	12.0-73.0
Age category -n (%)			
Adolescents, 12 - <18 years	5 (3.2)	4 (2.6)	9 (2.9)
18 - 65 years	128 (83.1)	126 (81.3)	254 (82.2)
>65 years	21 (13.6)	25 (16.1)	46 (14.9)
Sex -n (%)			
Female	62 (40.3)	64 (41.3)	126 (40.8)
Male	92 (59.7)	91 (58.7)	183 (59.2)
Race -n (%)			
White	111 (72.1)	102 (65.8)	213 (68.9)
Black or African American	0	1 (0.6)	1 (0.3)
Asian	19 (12.3)	29 (18.7)	48 (15.5)
Other	8 (5.2)	4 (2.6)	12 (3.9)
Unknown	16 (10.4)	19 (12.3)	35 (11.3)
Ethnicity -n (%)			
Hispanic/Latino	8 (5.2)	12 (7.7)	20 (6.5)
Not Hispanic/Latino	94 (61.0)	88 (56.8)	182 (58.9)
Not Reported	29 (18.8)	36 (23.2)	65 (21.0)
Unknown	23 (14.9)	19 (12.3)	42 (13.6)
Weight (kg)			
n	150	152	302
Mean (SD)	67.5 (14.04)	66.2 (14.78)	66.9 (14.41)
Median	67.7	66.2	67.0
Q1-Q3	58.0-78.0	54.6-74.5	55.6-75.6
Min - Max	28.5-97.0	32.9-115.5	28.5-115.5
Height (cm)			
n	148	144	292
Mean (SD)	169.7 (9.86)	170.0 (10.16)	169.9 (9.99)
Median	170.0	170.0	170.0
Q1-Q3	161.9-177.5	163.0-177.0	162.0-177.0
Min - Max	128.7-195.0	146.0-200.0	128.7-200.0
Body mass index (kg/m^2)			
n	146	142	288
Mean (SD)	23.4 (4.24)	22.7 (4.15)	23.1 (4.20)
Median	23.3	22.5	23.1
Q1-Q3	20.4-26.2	19.9-24.7	20.1-25.5
Min - Max	13.5-34.4	13.9-35.7	13.5-35.7

Body Mass Index (kg/m^2) = weight (kg) / (height (m)**2).

Underlying disease history/transplant and GvHD related history

Disease history by treatment (FAS)

Disease history	RUX N=154	BAT N=155	All Patients N=309
Primary diagnosis classification-n (%)			
Malignant-leukemia/MDS	129 (83.8)	121 (78.1)	250 (80.9)
Malignant-lymphoproliferative	18 (11.7)	26 (16.8)	44 (14.2)
Non-malignant	1 (0.6)	5 (3.2)	6 (1.9)
Other	6 (3.9)	3 (1.9)	9 (2.9)
Diagnosis of underlying malignant disease-n (%)			
Acute lymphoblastic leukemia (ALL)	25 (16.2)	16 (10.3)	41 (13.3)
Acute myelogenous leukemia (AML)	58 (37.7)	63 (40.6)	121 (39.2)
Chronic myelogenous leukemia (CML)	6 (3.9)	2 (1.3)	8 (2.6)
Excess blasts 2,developed from Fanconi syndrome	1 (0.6)	0	1 (0.3)
Hodgkin's lymphoma	6 (3.9)	2 (1.3)	8 (2.6)
Multiple myeloma	2 (1.3)	5 (3.2)	7 (2.3)
Myelodysplastic disorder	26 (16.9)	29 (18.7)	55 (17.8)
Non-Hodgkin's lymphoma	9 (5.8)	19 (12.3)	28 (9.1)
Other acute leukemia	4 (2.6)	3 (1.9)	7 (2.3)
Other leukemia	6 (3.9)	8 (5.2)	14 (4.5)
Other	4 (2.6)	0	4 (1.3)
Diagnosis of underlying non-malignant disease-n (%)			
Histiocytic disorders	0	1 (0.6)	1 (0.3)
Sickle cell disease	1 (0.6)	1 (0.6)	2 (0.6)
Other	0	3 (1.9)	3 (1.0)
Diagnosis of underlying disease Other -n (%)			
Blastic neoplasm of plasmacytoid dentritic cells	0	1 (0.6)	1 (0.3)
Multiple myeloma and secondary acute myeloid leukemia	0	1 (0.6)	1 (0.3)
Myelofibrosis	2 (1.3)	0	2 (0.6)
Myeloma	0	1 (0.6)	1 (0.3)
Myeloproliferative neoplasm	1 (0.6)	0	1 (0.3)
Post polycythemia vera myelofibrosis	1 (0.6)	0	1 (0.3)
Primary myelofibrosis	1 (0.6)	0	1 (0.3)
Septic granulomatosis	1 (0.6)	0	1 (0.3)
Time from diagnosis of underlying disease to screening (years)			
n	154	154	308
Mean (SD)	2.16 (3.195)	1.72 (2.170)	1.94 (2.735)
Median	1.04	0.86	0.94
Min-Max	0.2 - 25.7	0.2 - 15.1	0.2 - 25.7
CIBMTR risk assessment-n (%)			
Low	46 (29.9)	46 (29.7)	92 (29.8)
Intermediate	43 (27.9)	48 (31.0)	91 (29.4)
High	61 (39.6)	55 (35.5)	116 (37.5)
Unknown	4 (2.6)	6 (3.9)	10 (3.2)

Table 14.1-7.2 (Page 1 of 2)
Transplant related disease history by treatment
Full Analysis Set

Disease history	RUX N=154	BAT N=155	All Patients N=309
Conditioning regimen type-n (%)			
MYELOABLATIVE	85 (55.2)	65 (41.9)	150 (48.5)
NON-MYELOABLATIVE	31 (20.1)	41 (26.5)	72 (23.3)
REDUCED INTENSITY	38 (24.7)	49 (31.6)	87 (28.2)
Total HCT-specific co-morbidity index score-n (%)			
0	70 (45.5)	63 (40.6)	133 (43.0)
1	30 (19.5)	27 (17.4)	57 (18.4)
2	24 (15.6)	19 (12.3)	43 (13.9)
3	9 (5.8)	26 (16.8)	35 (11.3)
4	12 (7.8)	6 (3.9)	18 (5.8)
>=5	6 (3.9)	6 (3.9)	12 (3.9)
Missing	3 (1.9)	8 (5.2)	11 (3.6)

Table 14.1-7.2 (Page 2 of 2)
Transplant related disease history by treatment
Full Analysis Set

Disease history	RUX N=154	BAT N=155	All Patients N=309
Stem cell type-n (%)			
BONE MARROW	19 (12.3)	30 (19.4)	49 (15.9)
PERIPHERAL BLOOD	134 (87.0)	118 (76.1)	252 (81.6)
SINGLE CORD BLOOD	1 (0.6)	7 (4.5)	8 (2.6)
Cytomegalovirus status at time of transplant-n (%)			
NEGATIVE	73 (47.4)	68 (43.9)	141 (45.6)
POSITIVE	81 (52.6)	87 (56.1)	168 (54.4)

Table 14.1-7.3 (Page 1 of 3)
Transplant donor information by treatment
Full Analysis Set

Disease history	RUX N=154	BAT N=155	All Patients N=309
Number of Donors	157	158	315
Donor age (years)			
n	152	145	297
Mean (SD)	36.2 (12.80)	37.3 (12.86)	36.7 (12.82)
Median	35.0	36.0	35.0
Min-Max	11.0 - 68.0	14.0 - 73.0	11.0 - 73.0
Donor sex-n (%)			
Female	54 (34.4)	55 (34.8)	109 (34.6)
Male	102 (65.0)	99 (62.7)	201 (63.8)
Missing	1 (0.6)	4 (2.5)	5 (1.6)
HLA typing method-n (%)			
HIGH LEVEL DNA	105 (66.9)	89 (56.3)	194 (61.6)
LOCI A, B, C, DQ: LOW LEVEL DNA, LOCUS DRB1: HIGH LEVEL DNA	18 (11.5)	22 (13.9)	40 (12.7)
LOCI A, B, C: LOW LEVEL DNA, LOCUS DRB1: HIGH LEVEL DNA	1 (0.6)	5 (3.2)	6 (1.9)
LOCI A, B: LOW LEVEL DNA, LOCUS DRB1: HIGH LEVEL DNA	3 (1.9)	1 (0.6)	4 (1.3)
LOCI A, B: SEROLOGIC, LOCUS DRB1: HIGH LEVEL DNA	1 (0.6)	1 (0.6)	2 (0.6)
LOCI A, B: SEROLOGIC, LOCUS DRB1: LOW LEVEL DNA	1 (0.6)	1 (0.6)	2 (0.6)
LOW LEVEL DNA	14 (8.9)	11 (7.0)	25 (7.9)
SEROLOGIC	10 (6.4)	19 (12.0)	29 (9.2)
Missing	0	1 (0.6)	1 (0.3)
HLA match score-n (%)			
10/10	77 (49.0)	74 (46.8)	151 (47.9)
9/10	29 (18.5)	18 (11.4)	47 (14.9)
8/8	12 (7.6)	13 (8.2)	25 (7.9)

Percentages are based on the number of donors.

Baseline demographics were well-balanced between the two treatment arms

In the general population that had undergone allo-transplantation, the incidence of Grade II-IV aGvHD in adolescent patients aged 12-18 years approximates 5% of that observed in adults >18 years (data Center for International Blood and Marrow Transplant Research [CIBMTR] data – unpublished data). In the present study, the proportion of adolescents was 2.9%.

Few patients had other than underlying malignant disease (0.6% and 3.2%). The predominant malignant diagnoses (AML, ALL, MDS) were generally balanced between the treatment arms. In addition, the CIBMTR risk categories (a validated tool to categorize groups of patients undergoing allogeneic stem cell transplantation, intended for research purposes to stratify patients in broad disease risk categories for retrospective or prospective studies), were also balanced between the treatment arms.

Furthermore, the majority of conditioning regiments were myeloablative, however, a slight imbalance was noted, 55.2% of subjects in the RUX arm received myelo-ablative conditioning prior to transplant compared to 41.9% in the BAT arm. Consequently, fewer patients in ruxolitinib arm received the less intense forms of conditioning therapy (non-myelo-ablative: 20.1% and reduced intensity: 24.7%)

compared to BAT arm (non-myelo-ablative: 26.5% and reduced intensity: 31.6%). T-cell depletion was performed in 39 patients overall.

The transplant donor median age was 35 years, and the majority were male (balance between arms). The HLA match score was 10/0 or 9/10 in most cases.

Most patients (75.3%) entered the study with low hematopoietic cell transplantation (HCT)-comorbidity scores (between 0 and 2).

Table 14.1-7.4a (Page 1 of 3)
aGvHD disease history by treatment
Full Analysis Set

Disease history	RUX N=154	BAT N=155	All Patients N=309
Time from transplant to diagnosis of aGvHD grade>=2 (day)			
n	154	155	309
Mean (SD)	54.77 (62.813)	58.25 (58.680)	56.52 (60.701)
Median	33.50	34.00	34.00
Min-Max	3.0 - 380.0	6.0 - 412.0	3.0 - 412.0
Time from diagnosis of aGvHD grade>=2 to randomization (day)			
n	154	155	309
Mean (SD)	29.57 (43.612)	23.26 (31.392)	26.41 (38.046)
Median	14.00	14.00	14.00
Min-Max	2.0 - 332.0	1.0 - 293.0	1.0 - 332.0
aGvHD grade when diagnosis of grade >=2 -n (%)			
Grade II	68 (44.2)	74 (47.7)	142 (46.0)
Grade III	68 (44.2)	62 (40.0)	130 (42.1)
Grade IV	18 (11.7)	19 (12.3)	37 (12.0)
Steroid refractory aGvHD criteria met -n (%)			
Progression after at least 3 days	35 (22.7)	43 (27.7)	78 (25.2)
Failure to achieve a response after 7 days	72 (46.8)	63 (40.6)	135 (43.7)
Failure on steroid taper	47 (30.5)	49 (31.6)	96 (31.1)

Table 14.1-7.4a (Page 2 of 3)
aGvHD disease history by treatment
Full Analysis Set

Disease history	RUX N=154	BAT N=155	All Patients N=309
Prior aGvHD therapy -n (%)			
Steroid only	12 (7.8)	18 (11.6)	30 (9.7)
Steroid + CNI	77 (50.0)	76 (49.0)	153 (49.5)
Steroid + CNI + Other systemic aGvHD treatment	56 (36.4)	49 (31.6)	105 (34.0)
Steroid + CNI + only aGvHD prophylaxis	41 (26.6)	30 (19.4)	71 (23.0)
Steroid + CNI + only aGvHD treatment	8 (5.2)	12 (7.7)	20 (6.5)
Steroid + CNI + both aGvHD prophylaxis and treatment	7 (4.5)	7 (4.5)	14 (4.5)
Steroid + Other systemic aGvHD treatment	9 (5.8)	12 (7.7)	21 (6.8)
Steroid + only aGvHD prophylaxis	8 (5.2)	8 (5.2)	16 (5.2)
Steroid + only aGvHD treatment	1 (0.6)	4 (2.6)	5 (1.6)
Time from steroid refractory aGvHD to randomization (day)			
n	154	155	309
Mean (SD)	3.38 (5.940)	3.14 (5.130)	3.26 (5.541)
Median	1.00	1.00	1.00
Min-Max	0.0 - 47.0	0.0 - 24.0	0.0 - 47.0

Table 14.1-7.4a (Page 3 of 3)
aGvHD disease history by treatment
Full Analysis Set

Disease history	RUX N=154	BAT N=155	All Patients N=309
Time from diagnosis of aGvHD grade>=2 to steroid refractory (day)			
n	154	155	309
Mean (SD)	26.19 (43.159)	20.13 (30.831)	23.15 (37.547)
Median	11.00	10.00	10.00
Min-Max	1.0 - 331.0	1.0 - 293.0	1.0 - 331.0
Overall aGvHD grade at baseline -n (%)			
Grade 0	4 (2.6)	1 (0.6)	5 (1.6)
Grade I	2 (1.3)	0	2 (0.6)
Grade II	50 (32.5)	54 (34.8)	104 (33.7)
Grade III	68 (44.2)	68 (43.9)	136 (44.0)
Grade IV	30 (19.5)	32 (20.6)	62 (20.1)
aGvHD organ involvement at baseline -n(%)			
Skin	93 (60.4)	74 (47.7)	167 (54.0)
Liver	36 (23.4)	26 (16.8)	62 (20.1)
Upper GI	28 (18.2)	37 (23.9)	65 (21.0)
Lower GI	96 (62.3)	115 (74.2)	211 (68.3)
Missing	4 (2.6)	1 (0.6)	5 (1.6)
Steroid dose at randomization (mg)			
n	149	151	300
Mean (SD)	152.48 (112.555)	149.45 (90.704)	150.95 (101.983)
Median	150.00	150.00	150.00
Min-Max	20.0 - 1250.0	20.0 - 850.0	20.0 - 1250.0

Note: Corticosteroid dose (mg) = methylprednisolone dose (mg) × 1.25 + prednisone dose (mg).

The median time to conversion from grade 2 to 4 aGvHD to SR-aGvHD was 10 days (range: 1.0 to 331.0 days). The median time, from aGvHD grad ≥2 diagnosis and randomization was 14 days. The median time from diagnosis of SR-aGvHD to randomization was 1 day.

The number of patients randomized at baseline with grade II, III, and IV aGvHD was 104 (33.7%), 136 (44.0%), and 62 (20.1%), respectively.

Prior aGvHD therapy was steroids + Calcineurin inhibitor (CNI) in 50.0% in RUX arm and 49.0% in BAT arm. Second most common prior aGvHD treatment was steroids as aGvHD prophylaxis.

Prophylactic aGvHD treatments (systemic and/or locally active) were balanced between the two treatment arms. The most commonly used prior treatment for a GvHD was a combination of steroids and CNI and the most common reason for SR-aGvHD was failure to achieve a response after 7 days of treatment with steroids (46.8% in ruxolitinib arm and 40.6% in BAT arm).

Current medical conditions were (except for medical history consistent with study indication) mostly related to post allo-HSCT and symptoms of aGvHD and/or use of high-dose steroid complications. Most conditions were reported in, more or less, the same frequency in both study arms. However, CMV infections were reported more frequently among patients randomized to RUX arm than to BAT arm (CMV infection 23.4% vs 18.1%, CMV colitis 1.9% vs 0; CMV enteritis 1.3% vs 0.6%; and CMV viremia 1.3% vs 0). In addition, CMV test positive was reported in medical history for 3.9% in ruxolitinib arm and 9.7% in BAT arm.

The most common organs involved were skin (54.0%) and lower GI tract (68.3%). More patients in the RUX arm had symptoms of skin (60.4%) and liver (23.4%), than BAT arm (skin: 47.7% and liver: 16.8%). Upper GI symptoms and lower GI symptoms were less frequent in RUX arm (18.2% and 62.3%) than in BAT arm (23.9% and 74.2%).

Concomitant therapies

From randomization up to the data cut-off date, concomitant medications were taken by 98.7% and 100% patients in ruxolitinib and BAT arm, respectively. The overall profile of concomitant medications was, overall, similar between the two treatment arms. In addition to corticosteroids and CNIs, the

frequent concomitant medications also included agents for treatment of infections, gastric motility enhancers and electrolytes.

Up to the data cut-off, more patients in the BAT arm (53.5%) received additional systemic aGvHD treatments than ruxolitinib arm (22.1%) during the treatment period. The most frequent additional systemic aGvHD treatment in BAT arm was ruxolitinib (39.4%) where most patients were crossing over, and etanercept (6.5%). No single additional systemic aGvHD treatment was given to >10% patients in the ruxolitinib arm. The most common by PT in ruxolitinib arm was ECP (5.8%).

Prohibited medications

Prohibited medication was taken by 29 (19.1%) and 21 (14.0%) of subjects (at least one medication or non-drug therapy in the RUX arm and the BAT arm, respectively.

Transfusions

A total of 79.6% patients in ruxolitinib arm and 80.0% patients in BAT arm underwent transfusion after start of randomized study treatment, the most common being packed red blood cells (RBC). Furthermore, 87.8% of the patients underwent transfusion in the Crossover period.

Numbers analysed

Efficacy analyses used the Full Analysis Set (FAS), following the Intent-to-Treat (ITT) principle, comprising all patients to whom study treatment has been assigned by randomisation. All 309 randomised patients (except one patient who was randomised but did not sign the study consent) were included in the FAS: n=154 in the RUX arm and n=155 in the BAT arm.

Supportive analyses were performed using the Per Protocol Set comprised of patients compliant with the requirements of the study. Fifty-six (36.4%) patients in the RUX arm and 68 (43.9%) patients in the BAT arm were excluded from the PPS, mainly due to intake of prohibited medications (20.1% and 18.1%) and due to different aGvHD overall grades used for randomisation between CRF and IRT (18.2% and 17.4%).

Outcomes and estimation

The primary analysis comprised of analyses of the primary and key secondary endpoint and was performed in all patients who completed Day 56 or who discontinued from study participation earlier and had a cut-off date of 25-Jul-2019. Most endpoints were either updated at time of the second analysis or analysed only at the second analysis 17-Jan-2020.

Primary efficacy results

The primary endpoint, ORR at Day 28, assessed by local investigator (INV) using protocol advised criteria (according to Harris et al), was met (see table).

According to the SAP, a patient was not be considered a responder at Day 28 if any of the following events occurs; in case of missing aGvHD assessment at baseline or Day 28, no CR or PR at Day 28 or in case there was use of Additional systemic therapy for aGvHD prior to Day 28. It has been clarified that missing data in the primary analysis are presented in category ‘unknown’, while the category ‘missing visits’ actually refers to invalid aGvHD assessments at Day 28.

Overall response rate at Day 28 (FAS) by INV (primary analysis)

	RUX N = 154		BAT N = 155				
	n (%)	95% CI	n (%)	95% CI	Odds ratio (RUX/BAT)	95% CI	p-value
Overall response							
Responders							
Complete Response	53 (34.4)		30 (19.4)				
Partial Response	43 (27.9)		31 (20.0)				
Non-responders							
No Response	7 (4.5)		10 (6.5)				
Mixed response	10 (6.5)		17 (11.0)				
Progression	4 (2.6)		13 (8.4)				
Other *	1 (0.6)		7 (4.5)				
Unknown	36 (23.4)		47 (30.3)				
Death	15 (9.7)		22 (14.2)				
Early discontinuation	17 (11.0)		16 (10.3)				
Missing visits	4 (2.6)		9 (5.8)				
Overall Response Rate (ORR: CR+PR)	**96 (62.3)**	**(54.2, 70.0)**	**61 (39.4)**	**(31.6, 47.5)**	**2.64**	**(1.65, 4.22)**	**<.0001**

N: The total number of subjects in the treatment group. It is the denominator for percentage (%) calculation.
n: Number of patients who were at the corresponding category.
The 95% CI for the response rate was calculated using Clopper Pearson exact method.
One-sided p-value, odds ratio and 95% CI were calculated using stratified Cochrane-Mantel-Haenszel test.
*Other: patient with additional systemic therapies along with CR/PR per investigator assessment.

The number of patients who received a new systemic treatment for aGvHD prior to Day 28 was 10 (6.5%) in the ruxolitinib arm and 39 (25.2%) in the BAT arm. The requested sensitivity analysis used actual response instead of imputed non-response, for patients who received a new systemic treatment for aGvHD prior to Day 28. The number of responders in this sensitivity analysis was higher in the ruxolitinib arm (98 patients; 63.6%) compared to the BAT arm (69 patients; 44.5%); the difference was statistically significant (p=0.0005) with an odds ratio of 2.22 (95% CI: 1.40, 3.53), i.e., supportive for the primary endpoint result.

No response or progression were seen in 4,5% and 2,6% in the RUX arm and 6,5% and 8,4% in the BAT arm, respectively.

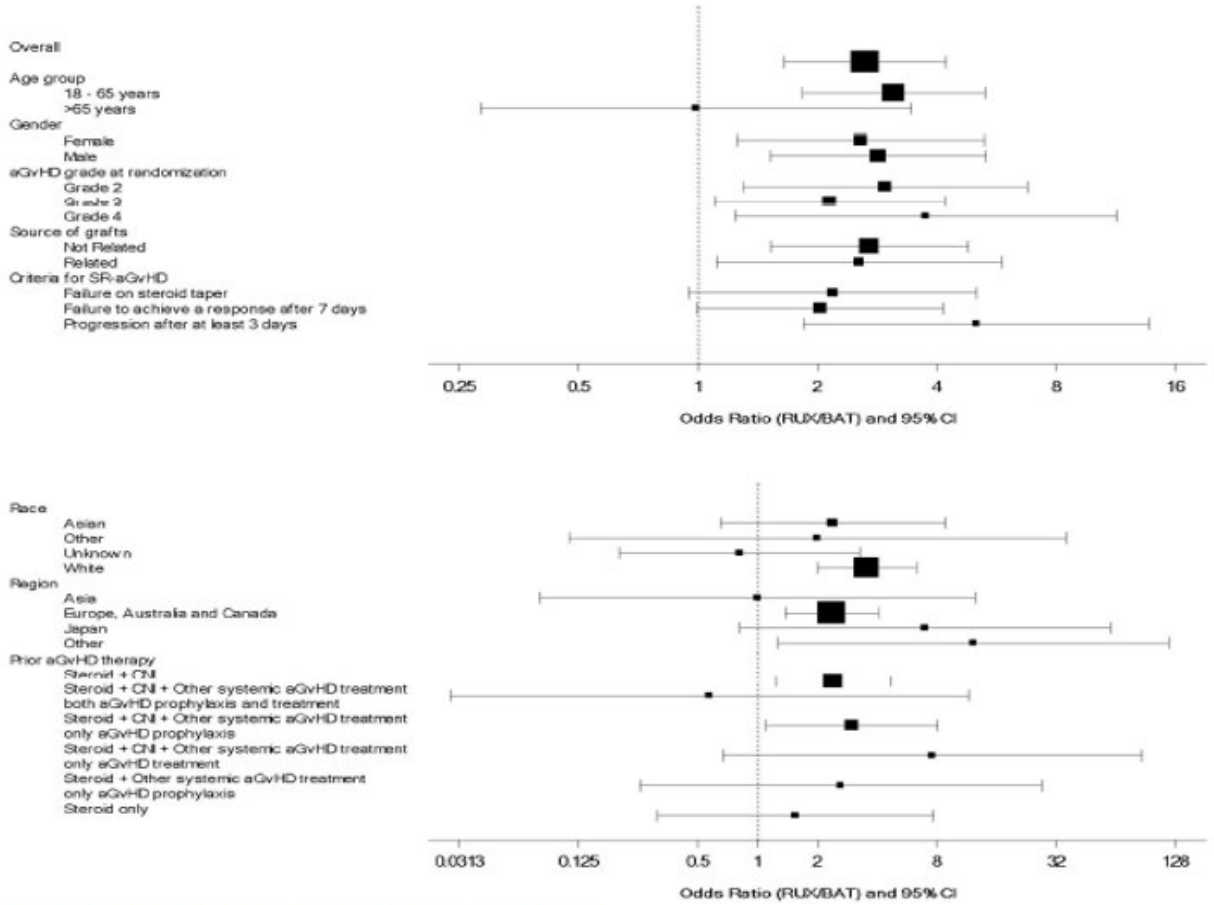
Unknow responses were 23,4% and 30,3% in the RUX arm and BAT arm respectively, in this category deaths accounted for 9,7% and 14,2%, early discontinuations for 11,0% and 10,3% and missing visits for 2,6% and 5,8% in respective RUX and BAT arm. With respect to the imbalance in numbers between the two arms, it is apparently more non-response imputations in the BAT arm than in the ruxolitinib arm, but further analyses were not specifically requested.

Subgroup analyses

Efficacy showed that the subgroup analyses (Day 28) were overall consistent with results for the primary endpoint, as demonstrated by the odds ratio (at Day 28) ranging from 1.55 to 12.28 in subgroups with evaluable data. Results favouring the BAT arm was only seen for patients >65 years, patients with unknown race and patients who received prior "steroid+CNI+Other" systemic aGvHD treatment. The results in these subgroups cannot, however, be definitively concluded due to the wide confidence intervals and limited sample size and wide confidence intervals.

The subgroup analysis demonstrated an advantage for the RUX arm compared to the BAT arm with an odds ratio (RUX/BAT) for ORR of 2.96 (Grade 2), 2.15 (Grade 3) and 3.76 (Grade 4), respectively.

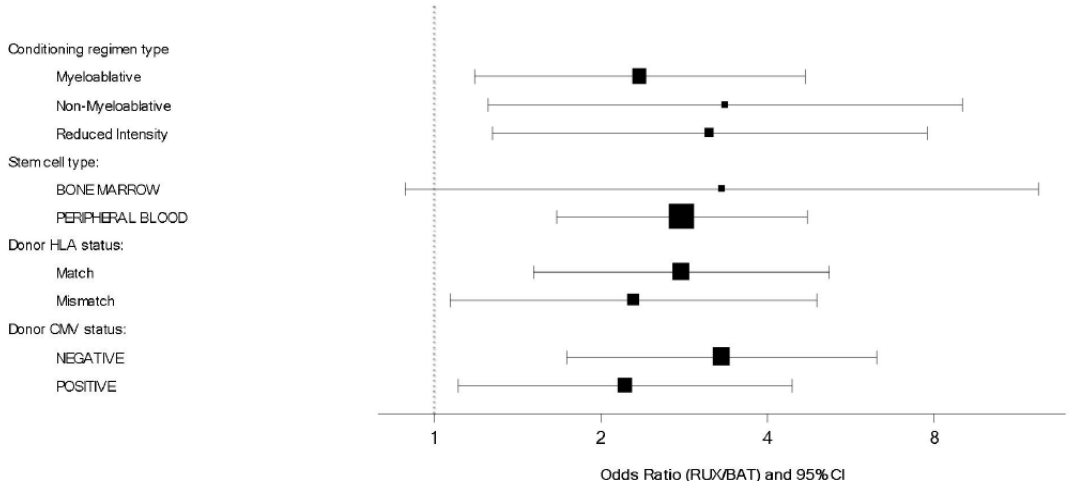
Forest plot of odds ratio with 95% confidence interval at Day 28 from <u>subgroup analysis</u> (FAS) -(A)

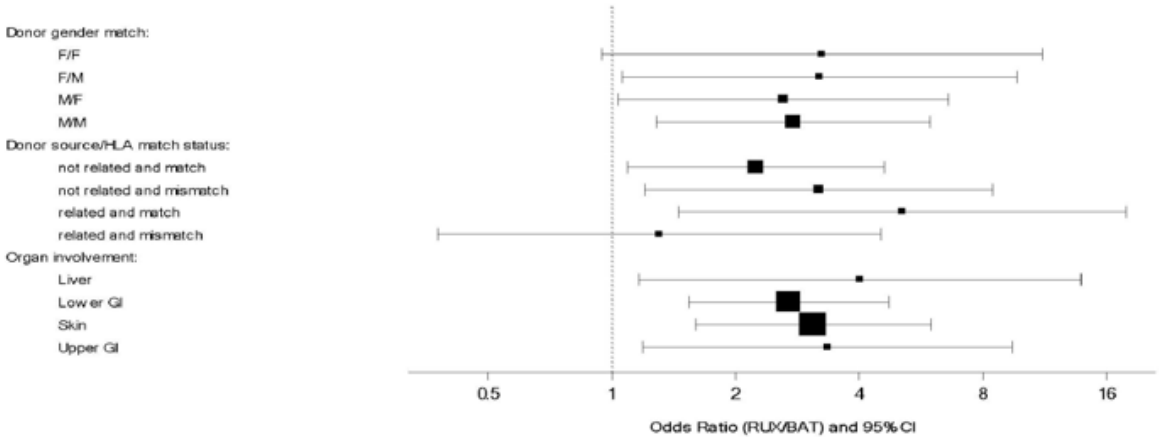


X-axis values are represented in natural log scale

Dotted lines shows no effect point, and (new) bold line shows overall treatment effect point (Odds ratio = 2.64). The area of the box indicates the weight of the sub group, measured by the size of subpopulation.

Forest plot of odds ratio with 95% confidence interval for ORR at Day 28 from subgroup analysis (FAS)-(B)





X-axis values are represented in natural log scale. Dotted lines shows no effect point.
The area of the box indicates the weight of the sub group, measured by the size of subpopulation.

Subgroup adolescents (12-18 years)

Table 14.2-1.1.4 (Page 1 of 29)
Overall response rate at Day 28 by subgroup
Full Analysis Set

Age group: Adolescents, 12 - <18 years

	RUX N= 5		BAT N= 4		Odds ratio	
	n (%)	95% CI	n (%)	95% CI	(RUX/BAT)	95% CI
Overall response						
Responders						
Complete Response (CR)	3 (60.0)		3 (75.0)			
Partial Response (PR)	1 (20.0)		0			
Non-responders						
No Response	0		0			
Mixed Response	0		0			
Progression	0		0			
Other *	0		0			
Unknown	1 (20.0)		1 (25.0)			
Death	0		0			
Early discontinuation	1 (20.0)		1 (25.0)			
Missing visits	0		0			
Overall Response Rate (ORR: CR+PR)	4 (80.0)	(28.4,99.5)	3 (75.0)	(19.4,99.4)	NE	(NE,NE)

The odds ratio in adolescent patients were inconclusive due to low number of evaluable patients (N=9).

A sensitivity analysis was performed using Fisher's exact test to compare ORR at Day 28 between the two treatment groups. The result was supportive of the primary analysis (odds ratio 2.55, with 95% CI (1.61, 4.03); p-value <0.0001) in favour of the RUX arm.

Supportive analysis for ORR at Day 28 (PPS population)

ORR in the PPS population showed similar results compared to the ORR rate in the FAS population (63.9% in the RUX arm vs 39.1% in the BAT arm in the PPS population and 62,3% in the RUX arm vs 39.4% in the BAT arm in the FAS population), indicating robustness of the results.

Table 14.2-1.1.2 (Page 1 of 1)
Overall response rate at Day 28
Per-Protocol Set

	RUX N = 97 n (%)	95% CI	BAT N = 87 n (%)	95% CI	Odds ratio (RUX/BAT)	95% CI	p-value
Overall response							
Responders							
Complete Response	36 (37.1)		14 (16.1)				
Partial Response	26 (26.8)		20 (23.0)				
Non-responders							
No Response	5 (5.2)		7 (8.0)				
Mixed response	7 (7.2)		8 (9.2)				
Progression	2 (2.1)		7 (8.0)				
Other *	0		4 (4.6)				
Unknown	21 (21.6)		27 (31.0)				
Death	9 (9.3)		13 (14.9)				
Early discontinuation	10 (10.3)		7 (8.0)				
Missing visits	2 (2.1)		7 (8.0)				
Overall Response Rate (ORR: CR+PR)	62 (63.9)	(53.5,73.4)	34 (39.1)	(28.8,50.1)	2.86	(1.55,5.27)	0.0005

Overall Response Rate (ORR) by initial Best Available Therapy (BAT) subgroups

Table 8. REACH-2: ORR at Day 28 by initial BAT

	Insufficient BAT compliance monitoring (worst case) per inspection report	N BAT (%)*	ORR BAT at day 28		N Rux	ORR RUX at day 28	
			ORR (%)**	95% CI		ORR (%)	95% CI
Total	16***	155	61 (39.4)	31.6–47.5	154	96 (62.3)	54.2–70.0
ATG	0	20 (12.9)	6 (30)	11.9-45.3			
Etanercept	2	22 (14.2)	10 (45.5)	24.4-67.8			
Everolimus	2	2 (1.3)	0	0-84.2			
ECP	0	41 (26.5)	18 (43.9)	28.5-60.3			
Infliximab	0	17 (11.0)	6 (35.3)	14.2-61.7			
Low-dose MTX	0	5 (3.2)	2 (40.0)	5.3-85.3			
MSC	0	15 (9.7)	9 (60.0)	32.3-83.7			
MMF	10***	25 (16.1)	8 (32.0)	14.9-53.5			
Sirolimus	2	3 (1.9)	2 (66.6)	9.4-99.2			

ATG anti-thymocyte globulin, ECP: extracorporeal photopheresis, MMF mycophenolate mofetil MSC mesenchymal stromal cells, MTX methotrexate, RUX: ruxolitinib, * uses all patients randomised to BAT as reference (denominator), **ORR within subgroup defined by initial BAT treatment, *** in REACH2 study, 2 patients flagged as having insufficient compliance monitoring in the inspector report were counted twice: Subjects 1 and 2 taking MMF were counted for both categories "administrated at home" and "administrated at home or hospitalized". Thus, the overall count is now 16 and not 18 as noted in inspector report

[1]Zeiser R, von Bubnoff N, Butler J, et al (2020) Ruxolitinib for Glucocorticoid-Refractory Acute Graft-versus-Host Disease. N Engl J Med; 382:1800-1810 - Supplementary material.

Shift in aGvHD by organ staging from baseline to Day 28

The trend of improvements favors the RUX arm in all assessed organs, although some baseline imbalances were detected, more patients in the RUX arm had acute GvHD involving skin (60.4%) and liver (23.4%), compared to the BAT arm (skin: 47.7% and liver: 16.1%). Organ staging data was missing for 23.4% patients in the ruxolitinib arm and 30.3% patients in BAT arm at Day 28, due to missing visits, discontinuation, or death.

Overall response rate at Crossover Day 28

Among the 49 patients who crossed over to RUX treatment between Day 28 and Week 24, the ORR at Crossover Day 28 was consistent with the primary analysis in the RUX arm: 67.3% (95% CI: 52.5, 80.1). The CR rate was higher compared with the primary analysis: 46.9%.

Secondary efficacy results

Durable ORR at Day 56 (key secondary endpoint)

Duration of response (DOR) is assessed for responders only and is defined as the time from first response until aGvHD progression or the date of additional systemic therapies for aGvHD. Onset of chronic GvHD, or death without prior observation of aGvHD progression were considered as competing risks.

Since the primary endpoint, ORR Day 28 was statistically significant at the primary analysis, the key secondary endpoint was also tested for significance at the primary analysis.

The drop in Durable ORR at day 56 compared to ORR Day 28 indicates that almost half of the responses, in both treatment arms, at Day 28 are no longer recorded at Day 56.

Durable overall response rate at Day 56 (FAS). Primary analysis.

	RUX N = 154		BAT N = 155				
	n (%)	95% CI	n (%)	95% CI	Odds ratio (RUX/BAT)	95% CI	p-value
Overall response							
Responders							
Complete Response	41 (26.6)		25 (16.1)				
Partial Response	20 (13.0)		9 (5.8)				
Non-responders							
No Response	1 (0.6)		1 (0.6)				
Mixed response	5 (3.2)		4 (2.6)				
Progression	0		0				
Other *	0		1 (0.6)				
Unknown	29 (18.8)		21 (13.5)				
Death	7 (4.5)		2 (1.3)				
Early discontinuation	13 (8.4)		15 (9.7)				
Missing visits	9 (5.8)		4 (2.6)				
Overall Response Rate (ORR: CR+PR)	**61 (39.6)**	**(31.8, 47.8)**	**34 (21.9)**	**(15.7, 29.3)**	**2.38**	**(1.43,3.94)**	**0.0005**

N: The total number of subjects in the treatment group. It is the denominator for percentage (%) calculation.
n: Number of patients who were at the corresponding category.
The 95% CI for the response rate was calculated using Clopper Pearson exact method.
One-sided P-value, odds ratio and 95% CI were calculated using stratified Cochrane-Mantel-Haenszel test.
Durable ORR at Day 56 is defined as the proportion of all patients who achieve a complete response (CR) or partial response (PR) at Day 28 and maintained a CR or PR at Day 56.
*Other: patient with additional systemic therapies along with CR/PR per investigator assessment.

Durable overall response rate at Crossover Day 56 (Crossover Analysis Set) (Second analysis)

Durable ORR at Crossover Day 56 was also tested for the crossover analysis set (the proportion of all crossover patients who achieved a CR or PR at Crossover Day 28 and maintained a CR or PR at Crossover Day 56). Among the 49 patients who crossed over, the durable ORR at Crossover Day 56

was 42.9% (95% CI: 28.8, 57.8). Thirty-six (73.5%) of these patients discontinued the cross-over treatment period with the most frequent reason for discontinuation being AEs (24.5%).

The durable ORR at Day 56 (FAS) during the randomized phase was consistent with the durable ORR for Crossover Day 56.

	RUX N= 49	
	n (%)	95% CI
Overall response		
Responders		
Complete Response (CR)	19 (38.8)	
Partial Response (PR)	2 (4.1)	
Non-responders		
No Response	1 (2.0)	
Mixed response	0	
Progression	1 (2.0)	
Other *	0	
Unknown	10 (20.4)	
Death	4 (8.2)	
Early discontinuation	4 (8.2)	
Missing visits	2 (4.1)	
Overall Response Rate (ORR: CR+PR)	21 (42.9)	(28.8,57.8)

N: The total number of patients in the treatment group. It is the denominator for percentage (%) calculation.
n: Number of patients who are at the corresponding category.
The 95% CI for the response rate is calculated using Clopper Pearson exact method.
Durable ORR at Crossover Day 56 is defined as the proportion of all crossover patients who achieve a complete response (CR) or partial response (PR) at Crossover Day 28 and maintain a CR or PR at Crossover Day 56.
*Other: patient with additional systemic therapies along with CR/PR per investigator assessment

- *Other secondary endpoints descriptive, not formally tested*

Overall response rate at Day 14

The ORR at Day 14 at the Primary analysis data cut-off was higher in RUX arm (63.0%; 95% CI: 54.8, 70.6) compared to BAT (47.1%; 95% CI: 39.0, 55.3; p<0.0029) with the OR for response in RUX arm compared to BAT arm being 1.98 and 95% CI: 1.24, 3.17.

Best overall response by Day 28

The BOR up to Day 28 at the Primary analysis data cut-off was higher in the RUX arm (81.8%) than in the BAT arm (60.6%; p<0.0001) with the odds ratio 3.07 and 95% CI: 1.80, 5.25 in favour of RUX.

The response rates at Day 14, 63% in the RUX arm and 39% in the BAT, compared to ORR Day 28 arm, 62.3% in the RUX arm and 39.4% indicate that the responses occur at an early timepoint after start of treatment. However, the BOR up to Day 28, 81.8% in the RUX arm and 60.6% in the BAT arm, also indicate that 19.5% and 21.2% of responses achieved up to Day 28, are no longer recorded. These "lost responses" are accounted for below, please refer to endpoint DoR and competing risks.

Duration of response

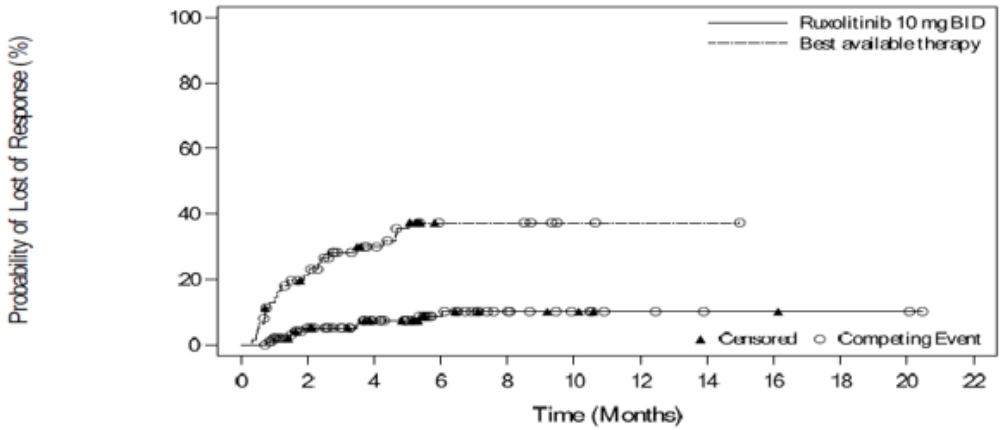
DoR was evaluated in patients who achieved a CR or PR at or before Day 28.

Duration of response (FAS set)

	RUX N=97	BAT N=62
Number of patients with events	9 (9.3)	22 (35.5)
Number of patients with competing risks	66 (68.0)	26 (41.9)
Death	34 (35.1)	14 (22.6)
Incidence of cGvHD	32 (33.0)	12 (19.4)
Number of patients censored	22 (22.7)	14 (22.6)
Duration of response (day)		
Median	163.0	101.0
Q1 - Q3	78.0 - 246.0	46.0 - 181.0
Min - Max	22.0 - 623.0	10.0 - 456.0
Estimated cumulative incidence and 95% CI at		
1 month	2.06 (0.39, 6.58)	12.97 (6.01, 22.66)
2 months	5.20 (1.92, 10.96)	21.38 (12.05, 32.47)
6 months	8.73 (4.03, 15.68)	37.34 (24.95, 49.71)
12 months	10.16 (4.91, 17.64)	37.34 (24.95, 49.71)
18 months	10.16 (4.91, 17.64)	NE (NE , NE)

N: number of patients whose overall response is CR or PR at Day 28.
The start date is the date of first documented response of CR or PR, which could be prior to or at Day 28.
The event is defined as the progression of aGvHD or addition of systemic therapies for aGvHD after Day 28.
The competing risks include death without prior observation of aGvHD progression and onset of cGvHD.
Duration of response is censored at the last response assessment.
Median and Quartiles are provided using Kaplan-Meier method.

Cumulative incidence curve of loss of response (FAS)



Number of patients still with response

Ruxolitinib 10 mg BID	97	76	53	29	17	11	6	3	3	2	2	0
Best available therapy	62	41	24	6	6	2	1	1	0	0	0	0

Cumulative number of competing events

Ruxolitinib 10 mg BID	0	14	30	44	53	58	61	64	64	64	64	66
Best available therapy	0	5	15	20	20	24	25	25	26	NA	NA	NA

The competing risks include death without prior observation of aGvHD progression and onset of cGvHD.
NA - Not Applicable

The median duration of response was longer in ruxolitinib arm (163 days, range: 22.0 to 623.0) compared to the BAT arm (101 days, range: 10.0 to 456.0).

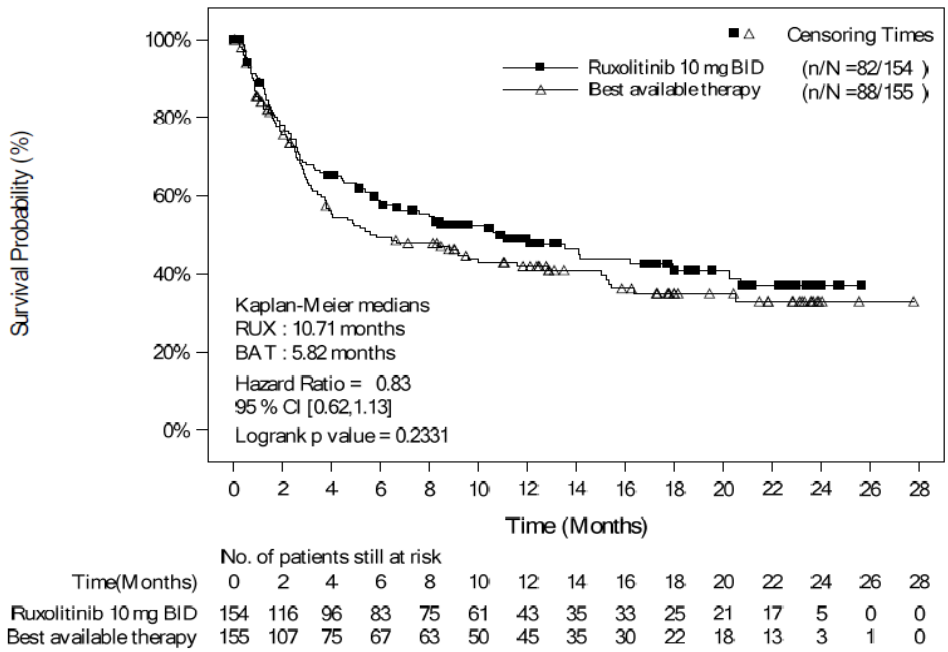
The probability of an event (progression or addition of systemic therapy for aGvHD) and 95% CIs at all estimated timepoints months, was lower in ruxolitinib arm compared to the BAT arm.

Overall survival

At the primary analysis, an interim OS analysis was performed with 72 (46.8%) deaths in the ruxolitinib arm and 79 (51.0%) deaths in the BAT arm. At the second analysis, there were 82 (53.2%) deaths in the ruxolitinib arm and 88 (56.8%) deaths in the BAT arm. OS median follow-up time was longer in ruxolitinib arm (7.34 months), compared to BAT arm (3.81 months).

Time-to-event analysis of OS stratified by aGvHD grade, suggested greater benefits of RUX for patients with lower baseline aGvHD grades, as shown by HRs closer to zero at lower grades. Patients with grade II aGvHD were shown to benefit most by RUX treatment (HR: 0.6; 95% CI: 0.34, 1.04). For patients with grade III the HR was 0.92, and in patients with grade IV the HR was 1.06.

Kaplan-Meier Curves of Overall survival (FAS)



No. of patients still at risk

Time(Months)	0	2	4	6	8	10	12	14	16	18	20	22	24	26	28
Ruxolitinib 10 mg BID	154	116	96	83	75	61	43	35	33	25	21	17	5	0	0
Best available therapy	155	107	75	67	63	50	45	35	30	22	18	13	3	1	0

p-value is obtained from the log-rank test.

Hazard ratio for overall survival (FAS)

	Event/N (%)	Log-rank Test Median Time (95% CI) (day)(2)	Cox Model Hazard Ratio (3)	95% CI
All Subjects (1)				
RUX	82/154 (53.2)	326 (182,547)	0.83	(0.62,1.13)
BAT	88/155 (56.8)	177 (115,392)		
Grade II				
RUX	22/ 53 (41.5)	547 (339,NE)	0.60	(0.34,1.04)
BAT	29/ 53 (54.7)	280 (116,NE)		
Grade III				
RUX	42/ 71 (59.2)	249 (85,617)	0.92	(0.60,1.41)
BAT	42/ 72 (58.3)	150 (87,458)		
Grade IV				
RUX	18/ 30 (60.0)	83 (50,NE)	1.06	(0.54,2.06)
BAT	17/ 30 (56.7)	197 (38,NE)		

(1) Cox PH model are stratified by aGvHD grade at randomization per IRT.
(2) Median (time to event) and its 95% CI are generated by KM estimation.
(3) Hazard Ratio of Rux versus BAT.
NE - Non Evaluable

There was no statistically significant impact on OS. The Kaplan-Meier (K-M) estimated median OS, was longer in RUX arm (10.71 months), compared to BAT arm (5.82 months). However, the cumulative incidence of OS event rates was similar, for the two treatment arms, over time, indicating no major difference in survival in the treatment arms. Longer follow-up of OS data up to the end of study may provide better understanding of the data.

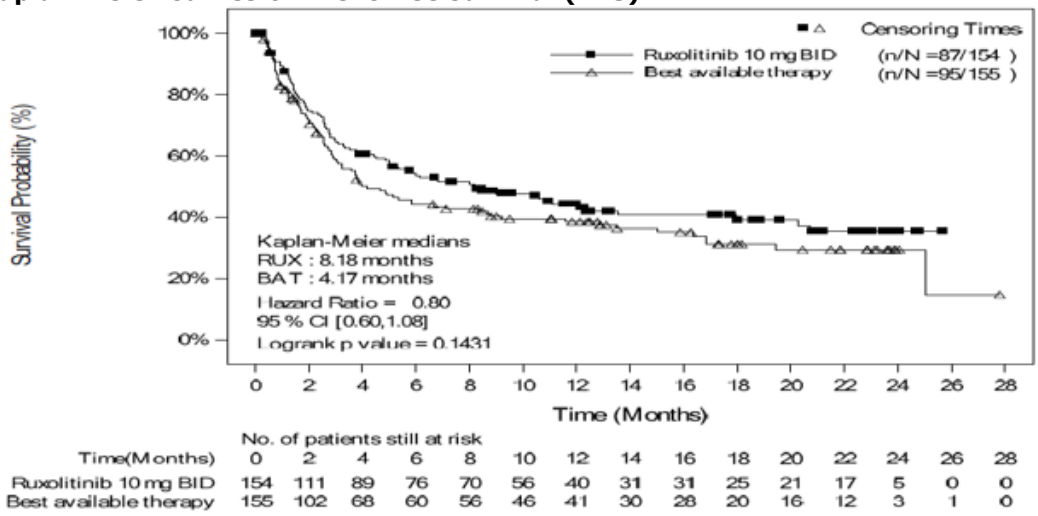
Given the study design, OS is primarily viewed as a safety endpoint. The allowance of cross-over creates a bias towards unity. In REACH-2 study, 49 (31.6%) out of 155 BAT subjects switched to ruxolitinib on or after Day 28. The rank-preserving structural failure time (RPSFT) model was applied post-hoc to investigate what would have been the OS analysis had patients not crossed over to ruxolitinib. After adjustment for cross-over with the RPSFT model, the HR estimate was 0.779 (95% CI: 0.515, 1.177) compared to 0.833 (95% CI: 0.616, 1.126) obtained in the pre-planned ITT analysis. Thus, the RPSFT results do not indicate any impact of cross-over on overall survival.

Event free survival

EFS was defined as the time from the date of randomization to the date of **i)** hematologic disease relapse/progression, **ii)** graft failure or death due to any cause. The EFS analysis at time of the second analysis cut-off date included 87 (56.5%) events in the RUX arm and 95 (61.3%) events in the BAT arm.

The K-M estimated median EFS was longer in RUX arm (8.18 months), compared to BAT arm (4.17 months). However, at DCO the difference in events between the arms have decreased and furthermore, a decreasing K-M probability estimates for survival at 2 months, 6 months, 12 months, 18 months and 24 months in both treatment arms, was seen.

Kaplan-Meier Curves of Event free survival (FAS)



No. of patients still at risk

Time(Months)	0	2	4	6	8	10	12	14	16	18	20	22	24	26	28
Ruxolitinib 10 mg BID	154	111	89	76	70	56	40	31	31	25	21	17	5	0	0
Best available therapy	155	102	68	60	56	46	41	30	28	20	16	12	3	1	0

p-value is obtained from the log-rank test.
The event includes hematologic disease relapse/progression, graft failure or death due to any cause.

Hazard ratio for event free survival (FAS)

		Log-rank Test	Cox Model	
	Event/N (%)	Median Time (95% CI) (day)(2)	Hazard Ratio (3)	95% CI
All Subjects (1)				
RUX	87/154(56.5)	249 (153,412)	0.80	(0.60,1.08)
BAT	95/155(61.3)	127 (92,254)		
Grade II				
RUX	25/53(47.2)	547 (207,NE)	0.63	(0.38,1.07)
BAT	32/53(60.4)	177 (94,512)		
Grade III				
RUX	44/71(62.0)	182 (79,412)	0.87	(0.57,1.32)
BAT	45/72(62.5)	118 (66,264)		
Grade IV				
RUX	18/30(60.0)	83(43,NE)	0.96	(0.50,1.85)
BAT	18/30(60.0)	92(37,592)		

(1) Cox PH model are stratified by aGvHD grade at randomization per IRT.
(2) Median (time to event) and its 95% CI are generated by KM estimation.
(3) Hazard Ratio of Rux versus BAT.
Event-free survival is defined as the time from the date of randomization to the date of hematologic disease relapse/progression, graft failure or death due to any cause.
NE - Non Evaluable

Sensitivity analysis for EFS including aGvHD progression as an event of interest produced similar results as primary analysis.

Time to event, EFS, stratified by GvHD grade, indicate a higher benefit for the RUX arm with lower aGvHD grade.

There reduction in risk of EFS event in the RUX arm relative to the BAT arm was <u>not</u> statistically significant and was also not corrected for multiplicity.

Failure free survival

FFS, which is a composite endpoint of (i) relapse or recurrence of underlying disease or death due of underlying disease, (ii) non-relapse mortality, or (iii) addition or initiation of another systemic therapy for aGvHD and the competing risk was cGvHD. Median FFS in the RUX arm was longer than in the BAT arm (4.86 months vs. 1.02 months; HR: 0.49, 95% CI: 0.37, 0.63; p<0.0001). The formal comparison using Log rank test was, however, not mentioned in the SAP and is considered only descriptively. As

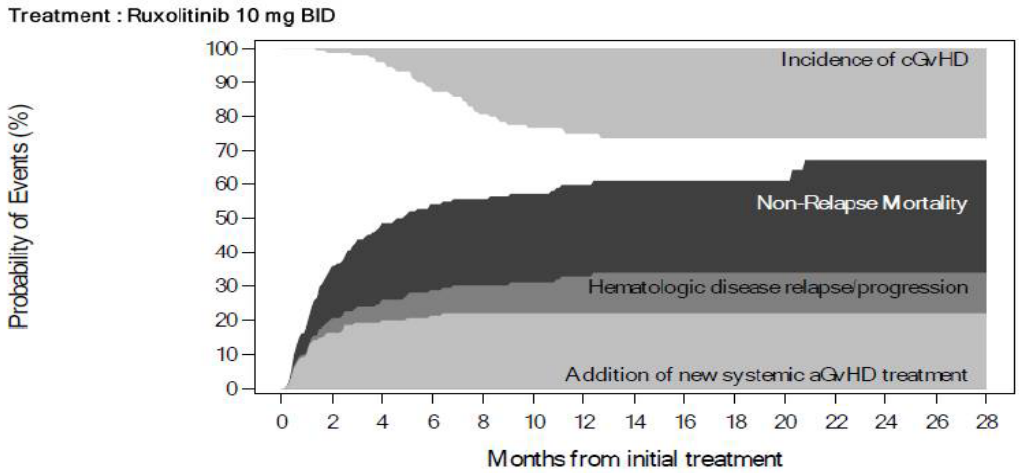
anticipated, FFS is driven by item (iii), while there is no indication of any impact of treatment on items (i) and (ii). FFS is not considered readily interpretable and these data are generated in an open-label study. Given these concerns, the FFS endpoint is not included in the product information.

Failure free survival (FAS)

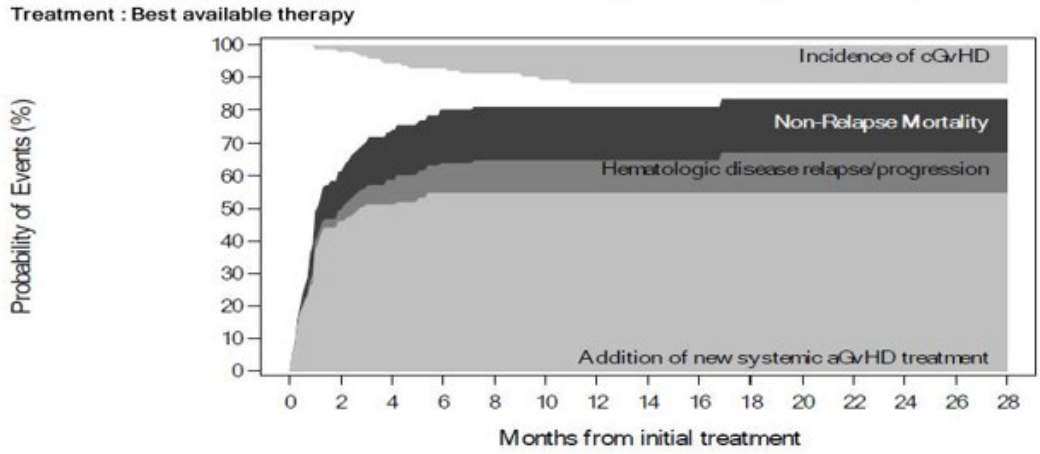
	RUX N=154	BAT N=155
Number of patients with events	91 (59.1)	121 (78.1)
Number of patients with competing risks	36 (23.4)	15 (9.7)
Incidence of cGvHD	36 (23.4)	15 (9.7)
Number of patients censored	27 (17.5)	19 (12.3)
Estimated cumulative incidence and 95% CI at		
1 month	18.47 (12.74, 25.04)	49.13 (40.94, 56.80)
2 months	35.82 (28.21, 43.48)	61.32 (53.00, 68.61)
6 months	54.07 (45.69, 61.71)	80.17 (72.52, 85.90)
12 months	59.59 (51.02, 67.14)	80.97 (73.37, 86.60)
18 months	60.76 (52.06, 68.38)	83.41 (74.17, 89.57)
24 months	NE (NE, NE)	83.41 (74.17, 89.57)

The competing risk includes onset of cGvHD.
Failure Free Survival includes hematologic disease relapse/progression, non-relapse mortality or addition of new systemic aGvHD treatment.
NE - Non Evaluable

Failure free survival by treatment (FAS)



Addition of new systemic aGvHD treatment	0%	17%	20%	21%	22%	22%	22%	22%	22%	22%	22%	22%	22%	22%	22%
Hematologic disease relapse/progression	0%	4%	6%	7%	8%	9%	11%	12%	12%	12%	12%	12%	12%	12%	12%
Non-Relapse Mortality	0%	15%	23%	25%	25%	26%	27%	27%	27%	27%	27%	33%	33%	33%	33%
Incidence of cGvHD	0%	1%	4%	13%	19%	23%	25%	26%	26%	26%	26%	26%	26%	26%	26%



Addition of new systemic aGvHD treatment	0%	46%	51%	55%	55%	55%	55%	55%	55%	55%	55%	55%	55%	55%	55%
Hematologic disease relapse/progression	0%	3%	8%	9%	10%	10%	10%	10%	10%	12%	12%	12%	12%	12%	12%
Non-Relapse Mortality	0%	12%	15%	16%	16%	16%	16%	16%	16%	16%	16%	16%	16%	16%	16%
Incidence of cGvHD	0%	2%	6%	7%	9%	11%	12%	12%	12%	12%	12%	12%	12%	12%	12%

The event includes hematologic disease relapse/progression, non-relapse mortality or addition of systemic aGvHD treatment.

Incidence of malignancy relapse/progression (updated at second analysis)

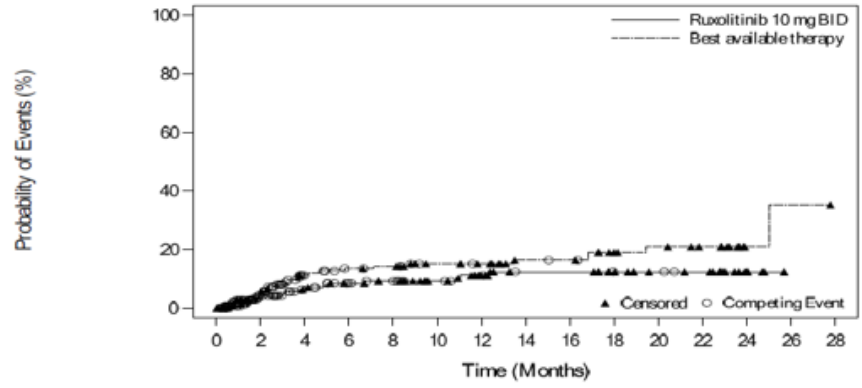
Out of 309 randomized patients, there were 147 subjects in each treatment arm with an underlying hematologic malignant disease at baseline.

The probability of malignancy relapse/progression was relatively low in both treatment arms, namely 16/147; 10.9% in ruxolitinib arm and 25/147; 17.0% in the BAT arm.

Incidence of Malignancy relapse/progression (FAS)

	RUX N=147	BAT N=147
Number of patients with events	16 (10.9)	25 (17.0)
Number of patients with competing risks	65 (44.2)	66 (44.9)
Death	65 (44.2)	66 (44.9)
Number of patients censored	66 (44.9)	56 (38.1)
Estimated cumulative incidence and 95% CI at		
1 month	0.69 (0.06, 3.51)	2.80 (0.92, 6.54)
2 months	4.21 (1.73, 8.46)	4.29 (1.75, 8.60)
6 months	8.49 (4.62, 13.83)	13.49 (8.32, 19.91)
12 months	11.12 (6.47, 17.17)	15.13 (9.61, 21.82)
18 months	12.40 (7.29, 18.93)	19.02 (12.36, 26.79)
24 months	12.40 (7.29, 18.93)	20.93 (13.56, 29.40)

N = number of patients with underlying hematologic malignant disease.
The competing risk includes death with non-relapse mortality for patients with underlying hematologic malignant disease.



Number of patients still at risk

Ruxolitinib 10 mg BID	147	108	86	74	68	51	38	30	30	24	20	17	5	0	0
Best available therapy	147	97	62	55	52	42	37	28	27	19	15	11	2	1	0

Cumulative number of competing events

Ruxolitinib 10 mg BID	0	28	45	52	54	58	60	62	62	63	63	65	65	65	NA
Best available therapy	0	33	55	59	60	63	64	64	65	66	66	66	66	66	66

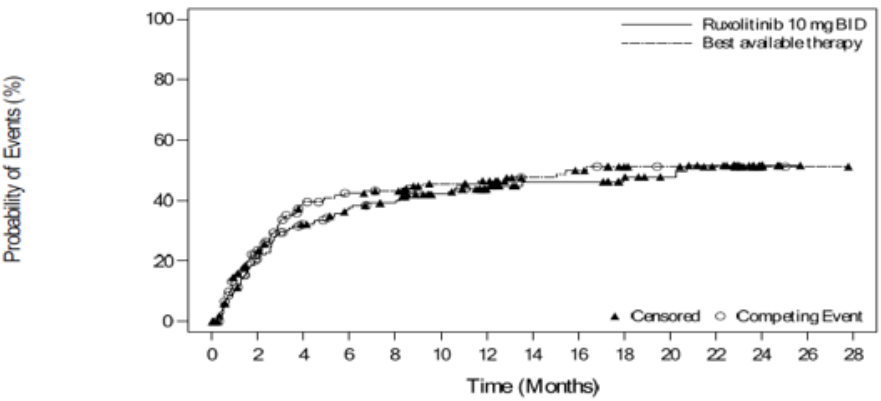
The competing risk includes death with non-relapse mortality.
NA - Not Applicable

Number of patients at risk is the number of patients who may develop malignancy relapse / progression at the given time point. Patients may experience event, competing risk, or become censored.

Non-relapse mortality

The analysis of NRM among all FAS patients included 69 (44.8%) patients with events in RUX arm and 70 (45.2%) patients in BAT arm. The cumulative incidence curves for NRM were overlapping for the RUX and BAT arms, indicating similar event rates over time. The competing risk (hematological disease relapse/progression) was low in both treatment arms (11.0% and 16.1%, respectively). However, censoring was high in both treatment arms (44.2% and 38.7%), implying a high proportion of patients who were alive and who had no relapse/progression of underlying malignancy.

Cumulative incidence curve of non-relapse mortality (FAS)



Number of patients still at risk

Ruxolitinib 10 mg BID	154	112	90	77	71	56	40	31	31	25	21	17	5	0	0
Best available therapy	155	103	68	61	57	47	42	31	28	20	16	12	3	1	0

Cumulative number of competing events

Ruxolitinib 10 mg BID	0	6	10	12	13	14	16	17	17	17	17	17	17	17	NA
Best available therapy	0	6	15	18	19	20	20	21	21	23	24	24	24	25	25

The competing risk includes hematologic malignancy relapse/progression. NA - Not Applicable

Death occurring after start of new systemic aGVHD therapy was not counted as an FFS event, if the patient already experienced an FFS event at the start of new systemic acute GvHD therapy. In other

words, NRM death was only counted as FFS event if it occurred before the other FFS events (hematologic disease relapse/progression and addition of new systemic aGvHD). For NRM as sole endpoint, all NRM deaths were counted as events, as start of new therapy would not interfere with this endpoint. Based on this, it is understood that the cumulative incidence of NRM is generally higher than the probability of NRM in the analysis of FFS, which was also the case in in this study. Moreover, NRM appears to be higher for ruxolitinib versus BAT in the FFS analysis because in the BAT arm many more patients had FFS events due to start of new systemic acute GvHD therapy than in the ruxolitinib arm.

Cumulative steroid dosing until Day 56

A successful tapering of steroids is a crucial step for the wellbeing of a patient after an allo-HSCT.

Tapering of steroids preceded any tapering of other immunosuppressants and should, according to the protocol, not be initiated before Day 7.

At Day 56, a numerically larger proportion of patients in RUX arm (21.4%; 95% CI: 15.2, 28.8) had tapered off corticosteroids than in BAT arm (14.8%; 95% CI: 9.6, 21.4) in the Primary analysis. Similar results were shown at the Secondary analysis (22.1% vs. 14.8%). The relative dose intensity and cumulative dose of steroids were similar between the RUX and BAT arms at all assessments.

Average weekly steroid dosing (FAS)

	Time RUX N=154	BAT N=155	Odds ratio
Completely tapered off by Day 56 - n (%) (95% CI)	34 (22.1) (15.8, 29.5)	23 (14.8) (9.6,21.4)	
			1.63 (0.91, 2.92)
≤50% RDI	45 (29.2) (22.2, 37.1)	37 (23.9) (17.4, 31.4)	
>50% RDI	106 (68.8) (60.9, 76.0)	117 (75.5) (67.9, 82.0)	

Dose intensity and relative dose intensity include days of zero dose in the calculation.
The planned dose intensity of steroid is the starting dose defined at the time of randomization per each patient.
Odds ratio and 95% confidence limits were calculated from Fisher's exact test.

A similar proportion of patients in both treatment arms were receiving ≤50% of RDI at Day 56 (29.2% and 23.9%).

Tapering of CNI and/or ruxolitinib would be initiated once the patient stopped corticosteroids. CNI taper was performed as per institutional guidelines.

Primary analysis: Duration of CNI:s from time of randomization

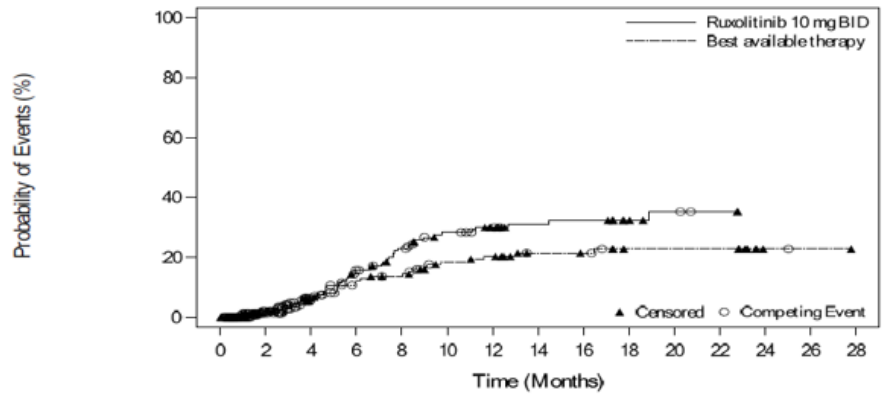
Table 14.3-2.10 (Page 1 of 1)
Duration of calcineurin inhibitors (CNIs)
Safety Set

	RUX N=152	BAT N=150
Total number of subjects -n (%)	128 (84.2)	122 (81.3)
Duration of treatment period (days)		
Mean (SD)	94.1 (74.69)	61.2 (53.17)
Median	73.5	32.5
Q1-Q3	37.5-142.0	26.0-90.0
Min - Max	1.0-396.0	1.0-199.0
Duration of treatment period categories -n (%)		
<= 28 days	24 (15.8)	49 (32.7)
>28 - 56 days	26 (17.1)	27 (18.0)
>56 - 112 days	36 (23.7)	23 (15.3)
>112 - 168 days	19 (12.5)	15 (10.0)
>168 - 336 days	21 (13.8)	8 (5.3)
>336 - 672 days	2 (1.3)	0
>672 days	0	0
Subject-time years	33.0	20.4

Table above, indicate a substantially longer use of CNI in the RUX arm. However, concomitant medications, such as CNI, were only captured in the study database up to 30 days after the last dose of the study treatment, thus including patients that that crossed over from BAT to ruxolitinib arm whose exposure to CNI was only collected until BAT discontinuation/the time of crossover. Hence, the reported shorter use of CNIs in the BAT arm reflects the study design rather than the actual exposure time to CNIs.

Incidence of cGvHD

Cumulative incidence of cGvHD (FAS)



Number of patients still at risk

Ruxolitinib 10 mg BID	154	110	82	57	43	28	20	13	12	7	4	2	0	0	0
Best available therapy	155	100	61	48	43	29	25	16	15	9	9	9	2	1	0

Cumulative number of competing events

Ruxolitinib 10 mg BID	0	37	58	66	68	72	75	77	77	77	77	79	79	NA	NA
Best available therapy	0	41	71	76	77	81	81	82	82	84	84	84	84	85	85

The competing risk includes deaths without prior onset of cGvHD and hematologic disease relapse/progression. NA - Not Applicable

A total of 45 (29.2%) patients in the RUX arm and 29 (18.7%) patients in BAT arm had developed cGvHD at DCO, i.e. more patients in the RUX arm had experienced an event of cGvHD. However, the difference in favor for the BAT arm, was seen mainly in Grade II aGvHD at randomization, while the cGVHD events were similar between the two treatment arms with respect to those with an aGvHD Grade 3 and 4 at randomization.

Competing events (deaths without prior onset of cGvHD and hematologic disease relapse/progression) were similar between the treatment arms.

The median time to onset of cGvHD was longer in the RUX arm (181.0 days) than the BAT arm (142.0 days). A majority of cGvHD events were reported mild at time of onset in both treatment arms.

Health related Quality of Life outcomes analysis

EQ-5D-5L questionnaire and FACT-BMT completion rates as a percentage of available patients were similar between the two arms throughout the study. Baseline scores were similar between both treatment arms. In both the randomized treatment and crossover periods, there was an overall improvement in all aspects of EQ-5D-5L and FACT-BMT questionnaires in both arms and no statistical analyses were performed.

Graft failure (Chimerism)

Occurrence of graft failure was reported as an event and also as an AE.

A small proportion of patients (9/309), 5 patients in ruxolitinib arm and 4 patients in BAT arm, experienced graft failure by the time of Second analysis cut-off date (no change in these numbers from the Primary analysis cut-off). In the ruxolitinib arm, 3/5 graft failures were detected by a drop in chimerism, and two patients discontinued due to graft loss. In the BAT arm, 1/4 graft failures were detected by a drop in chimerism, and one patient discontinued due to graft loss.

In the present study, the rate of graft failure is low and also comparable in the RUX arm and the BAT arm.

Supportive study

Title of Study: *A Single-Cohort, Phase 2 Study of Ruxolitinib in Combination With Corticosteroids for the Treatment of Steroid-Refractory Acute Graft Versus-Host Disease*

Study number: INCB 18424-271, REACH-1

Study period: This study was conducted at 26 study centres in the United States and the study period was 27 Dec 2016 to 05 Jun 2019 (final DCO).

Treatment for an individual participant could continue for **as long as benefit** was being observed and/or until treatment withdrawal criteria were met. The study ended, and the final analysis occurred when 75% of participants achieved 2-year NRM, died or were lost to follow-up.

Statistical methods

No formal statistical tests were performed. All CIs were 95%.

Outcomes/endpoints

All efficacy analyses were conducted in the efficacy evaluable population, defined as all patients enrolled in the study (N=71).

The primary objective was to evaluate the efficacy of ruxolitinib in combination with corticosteroids in participants with Grades II to IV SR-aGVHD as assessed by ORR at Day 28 (CR, VGPR or PR) based on CIBMTR response criteria and before the start of new anti-aGvHD therapy, if applicable.

Treatments

Patients received ruxolitinib at a starting dose of 5 mg b.i.d and the dose could be increased to 10 mg b.i.d if hematologic parameters were stable and no treatment-related toxicity was observed after the first 3 days of treatment. Dose-selection was based on publications (Spoerl et al, 2014 and Zeiser et al 2015), both using ruxolitinib as an add on treatment in SR-GvHD.

Participants received prednisone 2.5 mg/kg per day PO or methylprednisolone 2.0 mg/kg per day IV. Participants who previously began corticosteroid therapy at a different dose were permitted to remain on that dose if considered appropriate by the treating physician. Corticosteroids were tapered as per institutional guidelines at a rate commensurate with resolution of GVHD manifestations.

Study participants

Inclusion and exclusion criteria were similar to those for Study C2301. In the original protocol male or female, 18 years of age or older was an inclusion criterion, which was amended in Sept 2016, (Amend. 1) to include male or female, 12 years of age or older. No patients < 18 years of age were, however, included.

The majority of participants were < 65 years of age, with a median age of 58 years (range: 18-73 years) and 49.3% males and 50.7% females.

At baseline, mean donor chimerism was 95.3% (median: 100.0%; range 0%-100%).

Outcomes and estimation

Primary endpoint

Summary of Overall Response Rate at Day 28 (Efficacy Evaluable Participants)

Variable	Ruxolitinib (N = 71)
Number (%) of participants who had an overall response[a]	40 (56.3)
95% CI for ORR	(44.0, 68.1)
Responders	
CR	19 (26.8)
VGPR	6 (8.5)
PR	15 (21.1)
Nonresponders	
MR	3 (4.2)
NR	2 (2.8)
PD	2 (2.8)
Other	1 (1.4)
Missing[b]	23 (32.4)
Death	10 (14.1)
Discontinuation	12 (16.9)
Missing visits	1 (1.4)

[a] Participants who had a CR, VGPR, or PR at Day 28 response assessment or other response assessments within ± 2 days of Day 28, on or before the start of new anti-GVHD therapy (if applicable).
[b] Participants with missing assessment were considered nonresponders.

The 6-month DOR was the key secondary endpoint of the study. The median DOR for the 56.3% of patients who had response at Day 28 was 669.0 days, and for 76.1% of patients who had response at

any time point was 345.0 days. The 6-month event-free probability estimate for DOR based on response at Day 28 and response at any time point was 68.2% and 62.1%, respectively.

NRM: A total of 40 participants (56.3%) died from causes other than relapse of the underlying malignancy. The cumulative incidence rates at 6, 9, and 12 months were 44.3% (95% CI: 32.5, 55.5), 47.3% (95% CI: 35.2, 58.5), and 53.4% (95% CI: 40.9, 64.3), respectively.

Relapse rate: Five participants (7.0%) had a relapse of the underlying malignancy; the relapse had a fatal outcome in 4 of those participants (5.6%).

FFS: A total of 11 participants (15.5%) were still alive and had not had a relapse/progression of the underlying malignancy, had not required additional therapy for aGVHD, and had not demonstrated signs or symptoms of cGVHD. The median FFS time was 85.0 days (95% CI: 42.0, 158.0).

OS: A total of 44 participants (62.0%) had died, and 27 participants (38.0%) were censored at the last date known to be alive. The median OS time was 232.0 days (95% CI: 93.0, 675.0). At Months 3 and 6 the OS probability was >50%, and at Months 9 and 12, the OS probability was >40%.

2.5.4. Steroid refractory chronic GvHD: title of Study

Title of study: A phase III randomized open-label multi-center study of ruxolitinib versus best available therapy in patients with corticosteroid-refractory chronic graft vs. host disease after allogeneic stem cell transplantation.

Study number: CINC424D2301. Study identifier: **REACH3.**

Participating countries/regions and (number of sites): EU (177, including Norway and UK), Switzerland (5), Turkey (15), United States (34) Saudi Arabia (10), Australia (6) India (6), Israel (18), Japan (37), Jordan (3), Republic of Korea (4), Canada (6) and Russia (8).

Study period: Study initiation date: 29-Jun-2017 (first patient first visit)

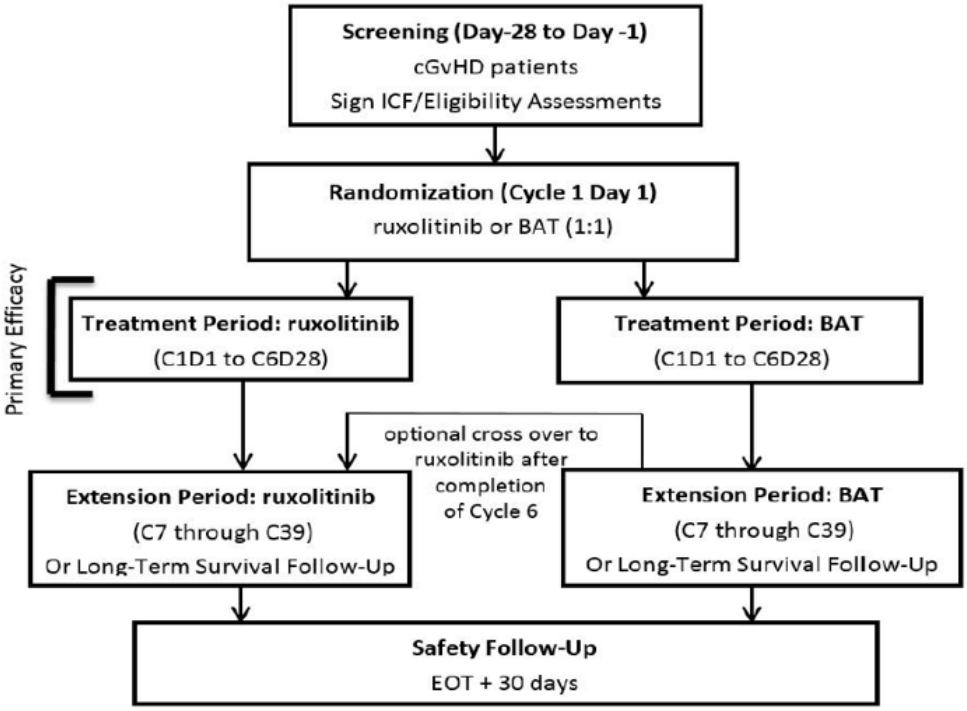
Data cut-off date: 08-May-2020 (data cut-off date for the primary analysis); study is ongoing.

The final analysis will occur once all patients have completed the study (up to 39 cycles/3 years from randomization). All available data from all patients up to end of study will be reported in a final CSR.

Phase of development (phase of this clinical study): Phase III

Study CINC424D2301 (study D2301, REACH3)) was performed to evaluate pharmacokinetics, efficacy and safety of ruxolitinib versus investigator-choice best available therapy (BAT) in adults and adolescents from 12 to less than 18 years old with moderate or severe SR-cGvHD following allogeneic HSCT.

Schematic Study design



Patients who crossed over to ruxolitinib treatment formed a stand-alone Cross-over analysis set for efficacy and safety assessment, not pooled with data from randomized ruxolitinib treatment

Methods

The present report presents the results of the primary analysis, which includes data up to the cut-off date of 08-May-2020, when all 329 patients randomized and completed the Cycle 7 Day 1 visit or discontinued from the study earlier.

At the time of that the COVID-19 pandemic occurred, most patients had completed Cycle 7 Day 1 visit or discontinued earlier. The impact of the pandemic on efficacy and safety assessment was reported as minimal.

Study participants

Inclusion criteria

- Male or female patients aged 12 or older at the time of screening informed consent.
- Had undergone alloSCT from any donor source (matched unrelated donor, sibling, haploidentical) using bone marrow, peripheral blood stem cells, or cord blood. Recipients of non-myeloablative, myeloablative, and reduced intensity conditioning were eligible.
- Clinically diagnosed cGvHD staging of moderate to severe according to NIH Consensus Criteria prior to Cycle 1 Day 1.
  - Moderate cGvHD: At least one organ (not lung) with a score of 2, 3 or more organs involved with a score of 1 in each organ, or lung score of 1
  - Severe cGvHD: at least 1 organ with a score of 3, or lung score of 2 or 3
- Evident myeloid and platelet engraftment (ANC >1,000/mm^3 and platelet count >25,000/mm^3)
- Patients received systemic or topical corticosteroids for the treatment of cGvHD for a duration of <12 months prior to Cycle 1 Day 1, and had a confirmed diagnosis of corticosteroid refractory cGvHD defined per 2014 NIH consensus criteria (Martin et al 2015) irrespective of the concomitant use of a calcineurin inhibitor, as follows:

  - A lack of response or disease progression after administration of minimum prednisone 1 mg/kg/day for at least 1 week (or equivalent)
    or
  - Disease persistence without improvement despite continued treatment with prednisone at >0.5 mg/kg/day or 1 mg/kg/every other day for at least 4 weeks (or equivalent)
    or
  - Increased prednisone dose to >0.25 mg/kg/day after two unsuccessful attempts to taper the dose (or equivalent) (criteria corresponding to steroid dependency)
- ECOG performance status of 0-2 OR Karnofsky Performance Score (KPS) OR Lansky Performance Score (LPS) of 60-100%.
- Patient accepted to be treated with only one of the following BAT options on Cycle 1 Day 1 (Additions and changes are allowed during the study, but only with BAT from the BAT options used in the study.)
  Note: Concomitant use of CNI and steroids was allowed. If any medication included in the BAT list was used as prophylaxis for underlying malignancy relapse, it was required to be discontinued prior to randomisation and entered into the eCRF. For patients randomised to either the RUX or BAT treatment arm, rituximab was allowed to be administered post-randomisation for the treatment of EBV. This EBV infection was captured either into the Medical History or Adverse Event eCRF. If any medication included in the BAT list was used as prophylaxis for cGvHD before entering this study, it was allowed to be continued post-randomisation and was captured on the Concomitant Medication eCRF.

Exclusion criteria

- Received two or more systemic treatments (BAT) for cGvHD in addition to corticosteroids ± CNI for cGvHD
- Patients that transition from active aGvHD to cGvHD without tapering off corticosteroids ± CNI and any systemic treatment
  - Patients receiving up to 30 mg by mouth once a day of hydrocortisone (i.e., physiologic replacement dose) of corticosteroids were allowed.
- Patients who were treated with prior JAK inhibitors for aGvHD; except when the patient achieved complete or partial response and had been off JAK inhibitor treatment for at least 8 weeks prior to Cycle 1 Day 1
- Failed prior alloSCT within the past 6 months from Cycle 1 Day 1
- Patients with relapsed primary malignancy, or who were treated for relapse after the alloSCT was performed.
- SR-cGvHD occurring after a non-scheduled DLI administered for pre-emptive treatment of malignancy recurrence.
- History of progressive multifocal leuko-encephalopathy (PML)
- Active uncontrolled bacterial, fungal, parasitic, or viral infection
- Patients on mechanical ventilation or had a resting O2 saturation <90% by pulse oximetry
- History or current diagnosis of cardiac disease indicating significant risk of safety for patients participating in the study such as uncontrolled or significant cardiac disease.
- Presence of severely impaired renal function defined by serum creatinine > 2 mg/dL (> 176.8µmol/L), renal dialysis requirement, or have estimated creatinine clearance <30 ml/min measured or calculated by Cockcroft Gault equation
- Cholestatic disorders, or unresolved sinusoidal obstructive syndrome/veno-occlusive disease of the liver (defined as persistent bilirubin abnormalities not attributable to aGvHD and ongoing organ dysfunction) or
  Total bilirubin >2mg/dL attributable to GvHD.

- Impairment of gastrointestinal (GI) function (unrelated to GvHD) or GI disease (unrelated to GvHD) that may significantly alter the absorption of oral Rux (e.g., ulcerative diseases, uncontrolled nausea, vomiting, diarrhea, malabsorption syndrome, or small bowel resection)
  Or
  Diarrhoea attributable to GvHD.
- Any corticosteroid therapy for indications other than cGvHD at doses >1 mg/kg/day methylprednisolone or equivalent within 7 days of Cycle 1 Day 1.
- Patient was receiving treatment with medications that interfere with coagulation or platelet function including, but not limited to, heparin or warfarin sodium (Coumadin®). Use of low molecular weight heparin is allowed. In patients in whom aspirin was indicated for secondary cardiovascular disease prevention, aspirin daily dose not exceeded 150 mg/day.
- Patient was receiving fluconazole at daily doses higher than 200 mg.
- Known allergies, hypersensitivity, or intolerance to systemic immunosuppressive therapy.
- Pregnant or nursing (lactating) women

Treatments

Patients were randomized 1:1 to receive either best available therapy (BAT, please see below) or ruxolitinib. The dose-selection is based on the same data as presented for study C2301 (aGvHD).

- Ruxolitinib 10 mg orally twice daily or
- Best available therapy (BAT) identified by the Investigator prior to patient randomisation among the following treatments: extracorporeal photopheresis (ECP), low-dose methotrexate (MTX), mycophenolate mofetil (MMF), mTOR inhibitors (everolimus or sirolimus), infliximab, rituximab, pentostatin, imatinib or ibrutinib. Addition or initiation of a new systemic therapy in the BAT arm was allowed only after documented lack of response intolerable toxicity, or cGvHD flare and was considered a treatment failure for both the primary and key secondary objectives.

Cycle length is 28 days.

Following randomization, change or addition of a new systemic therapy in the BAT arm due to documented lack of response or toxicity was allowed in the first 6 cycles, but was considered a treatment failure. At Cycle 7 Day 1 or later after randomization, patients randomized to BAT that did not achieve or maintain a CR or PR, or who developed toxicity to BAT were allowed to cross over from BAT to ruxolitinib.

The protocol defined BAT were based on a selection of therapies used as institutional standards since there is none uniformly used standard second line therapy in SR-cGvHD. Ibrutinib, included in the arsenal for the treatment of cGvHD after the protocol finalisation (however not an authorized indication in EU) was included as an of the electable choices of BAT, as a consequence of Amendment 1.

Initial BAT treatment D2301, REACH 3

Table 14.3-1.3 (Page 1 of 1)
Initial BAT treatment and number of BATs started before Cycle 7 Day 1
Safety set

	BAT N=158
Initial BAT treatment -n (%)	
Extracorporeal photopheresis (ECP)	55 (34.8)
Mycophenolate mofetil (MMF)	35 (22.2)
Ibrutinib	27 (17.1)
Low-dose Methotrexate (MTX)	10 (6.3)
Imatinib	8 (5.1)
Sirolimus	7 (4.4)
Rituximab	6 (3.8)
Everolimus	5 (3.2)
Infliximab	5 (3.2)
Number of BAT treatments started before Cycle 7 Day 1 -n (%)	
1	127 (80.4)
2	28 (17.7)
>2	3 (1.9)

CINC424D2301- Primary CSR

Tapering of concomitant immunosuppression

Tapering of immunosuppression therapy in responding patients was performed in two steps:

1. Taper of corticosteroids was attempted approximately 2 weeks after achieving a CR, based on protocol guidelines.

2. Taper of CNI and/or ruxolitinib was not initiated until the patient was off corticosteroids AND completed the assessments for Cycle 7 Day 1. Ruxolitinib was not be tapered prior to Cycle 7 Day 1 for patients initially randomized to the ruxolitinib arm.

Concomitant treatment

Supportive treatments per institutional guidelines for management of alloSCT patients with SRcGvHD were allowed. Patients may have received other cGvHD medications administered either as cGvHD prophylaxis or as treatment prior to randomization which could include CNI (cyclosporine or tacrolimus), and corticosteroids. Viral prophylaxis and antibiotics were allowed as needed for prevention and treatment of any infections.

Change or addition of new systemic immunosuppressive therapy after randomization and up to completion of Cycle 7 Day 1 was considered a treatment failure.

Ruxolitinib dose adjustments might be required particularly in patients treated with CYP450 modulators due to potential for drug-drug interactions leading to under- or over-exposure.

Permitted/prohibited medication according to the protocol are in line with Jakavi SmPC.

Cross-over

Patients randomised to BAT, and not achieving PR or better on or after Cycle 7 Day 1 were allowed to cross-over to the RUX treatment arm and were to follow the same treatment schedule as patients originally randomised to RUX treatment.

Duration of treatment

The End of Treatment (EOT) visit occurred when the patient permanently discontinued the study treatment and entered the Long-term survival follow-up, or completed the Randomized treatment period (ruxolitinib arm, or BAT patients not crossing over) or the Crossover period (BAT arm only). The total duration on study for an individual patient was 39 Cycles (156 weeks or 3 years).

Objectives

Primary objective

- To compare the efficacy of ruxolitinib vs. Investigator's choice Best Available Therapy (BAT) in patients with moderate or severe SR-cGvHD assessed by Overall Response Rate (ORR) at the Cycle 7 Day 1 visit.

Outcomes/endpoints

Endpoints according to Statistical Analysis Plan

Primary endpoint

Overall response rate (ORR) on Cycle 7 Day 1 after randomization, defined as the proportion of patients in each arm demonstrating a complete response (CR) or partial response (PR) without the requirement of additional systemic therapies for an earlier progression, mixed response or non-response. Scoring of response was relative to the organ score at randomization, using the established NIH consensus criteria.

Secondary endpoints

Key secondary endpoint-Composite time to event endpoint

-First key secondary endpoint defined as the time from date of randomization to the earliest of the following FFS events:

i) relapse or recurrence of underlying disease or death due to underlying disease **or**

ii) non-relapse mortality, **or**

iii) addition or initiation of another systemic therapy for cGvHD.

-The second key secondary objective was to assess the rate of patients with clinically relevant improvement of the modified Lee symptoms score at Cycle 7 Day 1 relative to baseline, as assessed by the rate of responders as per improvement ≥ 7 points of total symptom score (TSS) from baseline of the modified Lee Symptom Scale (mLSS).

Other secondary endpoints- please also refer to table in section 2.5.3.1.4

-Best overall response (BOR): Proportion of patients who achieved OR (CR+PR) at any time point (up Cycle 7 day 1 or the start of additional systemic therapy for cGvHD).

- Proportion of patients who achieved OR (CR+PR) at Cycle 4 Day 1.

- Duration of response (DOR) is assessed for responders only. DOR is defined as the time from first response until cGvHD progression, death, or the date of change/addition of systemic therapies for cGvHD.

- Overall survival (OS), defined as the time from the date of randomization to the date of death due to any cause.

- Non-relapse mortality (NRM), defined as the time from date of randomization to date of death not preceded by underlying disease relapse/recurrence.

-To assess proportion of patients with ≥50% reduction in daily corticosteroid dose at Cycle 7 Day1

-To assess proportion of patients successfully tapered off all corticosteroids at Cycle 7 Day 1

-To assess Malignancy Relapse/Recurrence (MR) is defined as the time from date of randomization to hematologic malignancy relapse/recurrence. Calculated for patients with underlying hematologic malignant disease.

- Change in FACT-BMT from baseline to each visit where measured. Change in EQ-5D from baseline to each visit where measured.

Efficacy analysis/ Criteria for evaluation

The severity of cGvHD at screening was assessed based on the number and degree of organ involvement according to the established NIH consensus criteria. For the definition of SR-CGvHD, please refer to inclusion criteria since the patient had to have a confirmed diagnosis of corticosteroid refractory cGvHD defined per 2014 NIH consensus criteria (Lee 2015).

The most frequently involved organs in patients with chronic GVHD are skin, mouth, and liver, with less frequent involvement of eye, lung, GI tract, joint/fascia, and genital tract.

Global and organ-specific cGvHD clinician assessments were performed at baseline, weekly for the first 4 weeks, and then every 28 days until Cycle 7 Day 1 according to the NIH Consensus Criteria described above. Following Cycle 7 Day 1, response was assessed on Cycle 9 Day 1 and every 12 weeks thereafter.

CR was defined as complete resolution of all signs and symptoms of cGvHD in all evaluable organs without additional therapies.

PR was defined as an improvement in at least one organ (e.g., improvement of 1 or more points on a 4 to 7-point scale, or an improvement of 2 or more points on a 10- to 12-point scale) without progression in other organs or sites, or addition/initiation of new systemic treatment.

Lack of response was defined as unchanged, mixed response, or progression.

PD was defined as worsening of at least one organ and no improvement (CR or PR) in any other organ

Mixed response was a CR or PR in at least 1 organ accompanied by progression in another organ

Unchanged response was defined as stable disease or absence of improvement in any organ involved by cGvHD

cGvHD Flare was defined as any increase in symptoms during taper of any immunosuppressive therapy for cGvHD, after an initial response (CR or PR). A cGvHD flare was not considered a treatment failure unless severity requires addition and/or change of another systemic immunosuppressive treatment.

cGvHD Recurrence was defined as the return of cGvHD symptoms after tapering off study treatment due to response. Following completion of a taper of systemic therapy, if worsening of cGvHD symptoms occur, the patient was allowed to resume treatment for cGvHD as per local institutional practice. However, this was to be documented as a recurrence of cGvHD.

Graft failure monitoring Patients will be monitored for any evidence of graft failure at each visit after Cycle 1 Day 1. In addition, considering that Graft failure is defined as initial whole blood or marrow donor chimerism >5% declining to <5% on subsequent measurements, donor chimerism will be also closely monitored. Donor chimerism after a hematopoietic stem cell transplant involves identifying the genetic profiles of the recipient and of the donor pre-transplant, and then evaluating the ratio of donor to recipient cells in the recipient's blood, or bone marrow.

Sample size

In a meta-analysis, Olivieri et al (2015) obtained an estimated pooled effect size for ORR for systemic treatment of SR-cGvHD of 0.66 (95% CI: 0.62–0.70). Sample size calculations were performed to achieve 90% power for different scenarios, assuming a targeted odds ratio of 2.35 and 2.5, respectively. A sample size of 324 patients was considered reasonable for this study based on stratified one-sided CMH test (alpha=2.5%, assuming same ORR and 50% of patients in each stratum).

With a sample size of 162 patients in each treatment (e.g., ORR=0.66 on Cycle 7 Day 1 for BAT for each stratum), an observed odds ratio greater than or equal to 1.68 is expected to result in statistically significant differences between the arms.

The sample size calculation seems appropriate. However, it is noted that the observed ORR rates are considerably lower than what was anticipated in the calculation. The ORR for Ruxolitinib estimated in the study is approximately 50%, i.e., only a fraction of the expected 82% as indicated above. The expected BAT effect (ORR of 66%) was not achieved in the study.

Randomisation

A stratified 1:1 randomisation was conducted, and patients received either ruxolitinib or BAT. They were stratified by cGvHD grade (Grade cGvHD moderate vs severe). Before randomization, BAT was selected by the Investigator.

Blinding (masking)

No blinding was performed due to the variety of treatments for SR cGvHD with different severity in the BAT arm.

Treatment phases/study conduct

-*Screening period (Day -28 to Day -1):* Prospective patients diagnosed with moderate or severe SR-cGvHD and who meet the inclusion and exclusion criteria will be consented prior to beginning screening assessments. The screening period can last 28 days (Day -28 to Day -1).

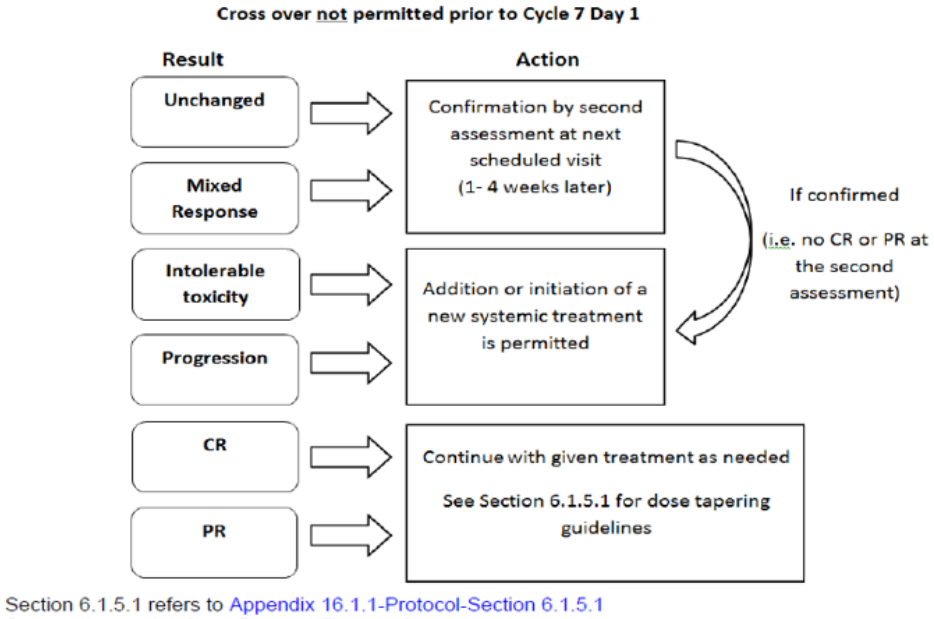
-*Main treatment period (Day 1 to EOT):* The total duration on study for an individual patient will be 39 Cycles (156 weeks or 3 years), inclusive of the randomized treatment period, cross over treatment period, and long-term survival follow-up.

-Crossover Treatment Phase: Patients randomized to the BAT arm who experience disease progression, have a mixed response, experience cGvHD flare, or experience toxicity to BAT are permitted to cross over to the ruxolitinib arm following completion of all assessments at Cycle 7 Day 1, or thereafter.

A maximum of 39 cycles of study treatment and follow-up is completed, inclusive of the randomized treatment period, cross over treatment period, and long-term survival follow up (i.e. patients who cross over on Cycle 9 will only complete visits until Cross-Over Cycle 30).

Primary efficacy period The Primary Efficacy Endpoint assessment will be conducted on Cycle 7 Day 1 of the randomized treatment period with a window of +/- 7 days.

Overall cGvHD Response assessment vs baseline (prior to Cycle 7 Day 1)



Extension period (Cycle 7 to Cycle 39)

-Safety follow-up (Last Dose + 30 days): All patients must have safety evaluations at least 30 days (+ 3 days) after the last dose of study treatment. During the safety follow up, adverse events, concomitant medications, transfusions, and monitoring of infections will be recorded. If an adverse event or a serious adverse event is detected, it should be followed until its resolution or until it is judged to be permanent.

-Long-term survival follow-up (EOT to 39 cycles on study): Patients who permanently discontinue the study treatment prior to the completion of 39 cycles on study for reasons other than achieving a CR or PR will enter the Long-Term Survival Follow-Up, and may be treated per Institutional practice. They will be followed approximately every 3 months by telephone call for survival, reporting of new cGvHD therapies until 39 cycles on study is completed, inclusive of randomized treatment, cross over treatment (BAT patients only), and long-term survival follow up.

Statistical methods

The primary efficacy variable, ORR at Cycle 7 Day 1, was planned to be analysed at the time when all patients have completed their Cycle 7 Day 1 visit or discontinued earlier. The data cut-off date for the primary analysis was 08-May-2020. An interim analysis on the primary endpoint ORR and key secondary endpoints FFS and mLSS, based on cut-off date 9-Jul-2019, was performed in October 2019.

The final analysis will occur once all patients have completed the study (up to 39 cycles/3 years from randomization).

Analysis sets

The Full Analysis Set (FAS) comprises all patients to whom study treatment has been assigned by randomization. According to the intent to treat principle, patients were analyzed according to the treatment and strata they have been assigned to during the randomization procedure.

The Per-Protocol Set (PPS) consists of a subset of the patients in the FAS who are compliant with requirements of the clinical study protocol (CSP).

The Safety Set includes all patients who received at least one dose of study treatment. Patients were analysed according to the study treatment received, where treatment received was defined as the randomized treatment if the patient took at least one dose of that treatment, or the first treatment received if the randomized treatment was never received.

The Cross-over Analysis Set (CAS) comprises all patients randomized to and who receive BAT, who then crossed over and received at least one dose of ruxolitinib. This analysis set was used for all analyses for crossover patients.

Primary analysis

The primary endpoint of the study was ORR at Cycle 7 Day 1 defined as the proportion of patients with CR or PR, according to the NIH Consensus Criteria. Note that response is relative to the assessment of cGvHD at randomization. A patient was not considered a responder if any of the following events occurs prior to the Cycle 7 Day 1 visit:

- Missing overall cGvHD response assessment at Cycle 7 Day 1
- No CR or PR at Cycle 7 Day 1
- Addition of or start of new systemic therapy for cGvHD.

The statistical hypotheses and method of analysis was similar to the primary analysis of the REACH 2 study. The Cochrane-Mantel-Haenszel (CMH) chi-square test, stratified by the randomization stratification factor (i.e., cGvHD moderate vs. severe), was used to compare ORR between the two treatment groups. The primary analysis was performed on the FAS according to ITT principle. One-sided p-value, odds ratio (OR) and 95% Wald confidence limits calculated from the stratified CMH test were presented. ORR was also summarized using descriptive statistics (N, %) by treatment arm along with two-sided exact binomial 95% CIs [Clopper and Pearson 1934].

Supportive analyses included re-run of the primary efficacy analysis on the PPS, and detailed descriptions of response rates and organ specific responses at Cycle 7 Day 1.

Subgroup analysis was performed for the primary efficacy endpoint to assess the homogeneity of treatment effect across demographic and baseline disease characteristics. Efficacy analyses in subgroups were purely exploratory and intended to explore the consistency of treatment effect.

ORR at Crossover Cycle 7 Day 1 was defined as the proportion of crossover patients with CR or PR at Crossover Cycle 7 Day 1. ORR was summarized using descriptive statistics (N, %) along with two-sided exact binomial 95% CIs [Clopper and Pearson 1934] using local investigators' overall response assessed at the Crossover Cycle 7 Day 1 visit and taking into account initiation or addition of new systemic therapy before this time point. Note that response was relative to the last assessment of cGvHD prior to or at the start date of crossover treatment (ruxolitinib). A patient was not considered a responder at Crossover Cycle 7 Day 1 if any of the following events occurs prior to Crossover Cycle 7 Day 1:

- Missing overall cGvHD response assessment at crossover or Crossover Cycle 7 Day 1
- No CR or PR at Crossover Cycle 7 Day 1
- Addition of or start of new systemic therapy for cGvHD.

Analysis of the key secondary endpoints

The first key secondary endpoint was Failure free survival (FFS), in line with the CHMP/PMDA's recommendation and defined as the time from date of randomization to the earliest of: i) relapse or recurrence of underlying disease or death due to underlying disease, ii) non-relapse mortality, or iii) addition or initiation of another systemic therapy for cGvHD. If a patient did not experience any of these events, FFS were censored at the latest contact data (on or before the cut-off date). A stratified

one-sided log-rank test was used to test if FFS hazard ratio (ruxolitinib arm versus BAT arm) was lower than 1. The stratification was based on the randomization stratification factors (i.e. cGvHD moderate vs severe).

No data imputation was performed for key secondary endpoints when missing values, censoring, or discontinuations occurred.

The second key secondary objective was to assess the improvement of symptoms based on the total symptom score (TSS) using the modified Lee Symptom Scale. A responder was defined as having achieved a clinically relevant reduction by ≥7 scores from baseline of the TSS. A patient was not considered a responder at Cycle 7 Day 1 if any of the following events occurs prior to or at the Cycle 7 Day 1 visit:

- Missing or insufficient data to calculate TSS at baseline or Cycle 7 Day 1
- No clinically relevant reduction of TSS at Cycle 7 Day 1
- Additional systemic therapy for cGvHD.

The Cochrane-Mantel-Haenszel chi-square test, stratified by the randomization stratification factor (i.e., cGvHD moderate vs severe), was used to compare the response rates between the two treatment groups.

Handling of missing values/censoring/discontinuations

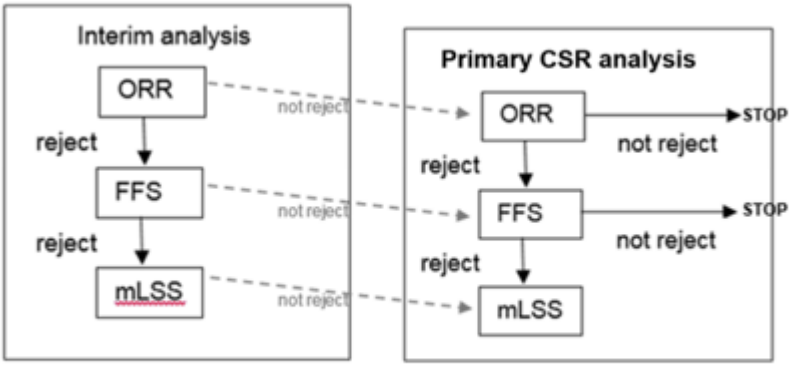
Patients with missing assessments that prevent the evaluation of the primary endpoint were considered non-responders on that treatment arm. This includes missing overall cGvHD response assessments at baseline and Cycle 7 Day 1. The same rule was applied for the evaluation of the ORR at Cycle 4 Day 1 and BOR, including missing cGvHD assessments at Cycle 4 Day 1.

No data imputation was to be applied. Addition or initiation of a new systemic therapy before Cycle 7 Day 1 in any arm was considered a treatment failure, and patients were counted as non-responder in the primary analysis.

Multiplicity

A fixed sequence hierarchical testing strategy has initially been planned for the primary and the two key secondary endpoints, with all tests being one-sided with significance level alpha=2.5%. With protocol amendment 1, an efficacy and safety interim analysis (IA) was added based on a 2-look group-sequential study design, and a hierarchical testing procedure was implemented to control the familywise error rate (FWER) at the one-sided 2.5% level of significance for the primary and the two key secondary endpoints at the IA and the primary CSR analysis (see figure below). (In line with the CHMP/PMDA's recommendation, FFS was used as the first key secondary endpoint for all regions except the US. A different testing sequence was used for the US (FDA), where mLSS was tested before testing FFS.)

Figure 8. Hierarchical testing strategy used for all regions except for the US (RoW)



ORR: overall response rate; FFS: Failure-free survival; mLSS refers to responder rate based on the modified Lee symptom score (mLss)
RoW: Rest of the world

The first efficacy look was planned when 194 (60% of the targeted 324 patients) had completed the Cycle 7 Day 1 visit or discontinued from the study earlier and data of assessments are available.

Interim analysis was performed based on efficacy data of the first 196 patients (60.5% of the targeted 324 patients) on cut-off date 9-Jul-2019. The resulting significance level alpha was 0.01176 for the interim analysis. With N=329 for the primary analysis the targeted alpha to re-test the hypotheses (if not rejected at the IA) was 0.01858. Following the overall hierarchical testing procedure, the hypotheses that were not rejected at the IA are to be tested again in this this CSR analyses spending the remaining alpha, i.e. IA results have to be taken into account when interpreting CSR results.

Interim analysis

One early safety interim analysis and an interim efficacy and safety analysis (both added in protocol amendment 1) were planned. An early safety analysis was generated when safety data on the first 80 randomized patients who have completed Cycle 4 Day 1 were available. The interim analysis was to be performed by an independent statistician and programmer when 194 (60% of the targeted 324 patients) have completed the Cycle 7 Day 1 visit or discontinued from the study earlier and data of assessments were available. Following a 2-look group sequential design a rho-spending function with parameter rho=1.5 was used as alpha spending function, for which operational characteristics were given in the amended protocol (dated 20-Dec-2017). If the number of patients at the IA is not exactly 194 at the time of the interim cut-off, the efficacy stopping bound would be recalculated based on the pre-specified alpha spending function.

The interim analysis was performed based on efficacy data of the first 196 patients (60.5% of the targeted 324 patients) on the significance level alpha of 0.01176. With N=329 for the primary analysis the targeted alpha to re-test the hypotheses (if not rejected at the IA) was 0.01858. The interim analysis p-values for the primary and the key secondary endpoints were: 0.0003, <0.0001 and 0.0151. At the CSR primary analysis, the corresponding p-values were <0.0001, <0.0001 and 0.0011. The 3rd endpoint (proportion of responders based on the TSS) was according to the applied hierarchical testing procedure not significant at the interim analysis but could be re-tested at the primary CSR analysis on 0.01858 alpha level and therefore was statistically significant.

Changes from the planned analyses

No relevant changes were made between the (amended) protocol statistical section and the SAP.

However, an error was detected in the programmed derivation of some of the mLSS subscales caused by an incorrect description in the SAP Amendment 2. The corrected description of the scoring is

documented in the SAP Addendum to Amendment 2. The programs were updated for the respective analysis data and all outputs for mLSS were re-produced. The impact of these updates on the overall results for mLSS was low, the key secondary endpoint remained statistically significant.

The addition of an interim efficacy and safety analysis in the protocol amendment 1 was made early during the study conduct (at the time when 15 subjects were screened and randomized) and is not likely to have been influenced by any observed study data. The analyses were performed in accordance with the amended protocol.

Of note, for the interpretation of the phase III study results, the convention is to use 2-sided tests and p-values. For that reason, 2-sided p-values are included in the SmPC.

Protocol amendments

The original protocol was dated 14-MAR-2017 and an amended protocol version 01 was effective from 20-DEC-2017; issued after 15 patients had been randomized.

The main purpose of the amendment was to extend the available patient population. Due to protracted enrolment and feedback from health authorities and investigators, several inclusion and exclusion criteria were revised. Additionally, following the approval of ibrutinib, the original protocol-defined list of BAT options would not allow for a complete comparative assessment versus ruxolitinib. Therefore, adding ibrutinib to the list of eligible BAT was necessary to reflect the complete list of treatment options for this patient population. Lastly, a Data Monitoring Committee was been added based on health authority feedback to include an independent review group.

At the time of the amendment finalization, 15 patients had been screened and all 15 had been randomized, and none of these patients had reached the 6-month visit that was used for the primary endpoint.

With regards to amendments concerning inclusion and exclusion criteria, please also refer to section 2.4.5.1.1.

There was one protocol amendment, issued 20-DEC-2017, after 15 patients had been randomized in study D2301, REACH3. There were several changes including revision of some inclusion and exclusion criteria as well as some secondary endpoints. Most of the changes being of clarifying nature, however, some of more clinically relevant character. The updated number of patients for extensive PK sampling is acknowledged as well as the important update of the window for patients receiving systemic or topical corticosteroids for the treatment of cGvHD, which was extended from six months to twelve months. Furthermore, patients who had received two or more systemic therapies for cGvHD were excluded, thus only patients who had received one prior replacement dose range was allowed. The amendment allowing a broader inclusion population, i.e. updated definition of overlap syndrome might have facilitated the inclusion rate.

With regard to the added DMC, interim analysis and the SAP addendum, these are endorsed. The addition of DMC and interim efficacy and safety analysis in the protocol amendment 1 was made early during the study conduct and is not likely to have been influenced by any observed study data. The scoring for mLSS was corrected prior to the CSR, and the key secondary endpoint remained statistically significant.

Results

This clinical study report (CSR) presents the results of the interim analysis and the primary analysis, which includes data up to the cut-off date of 08-May-2020, when all 329 patients randomized and completed the Cycle 7 Day 1 visit or discontinued from the study earlier.

Recruitment

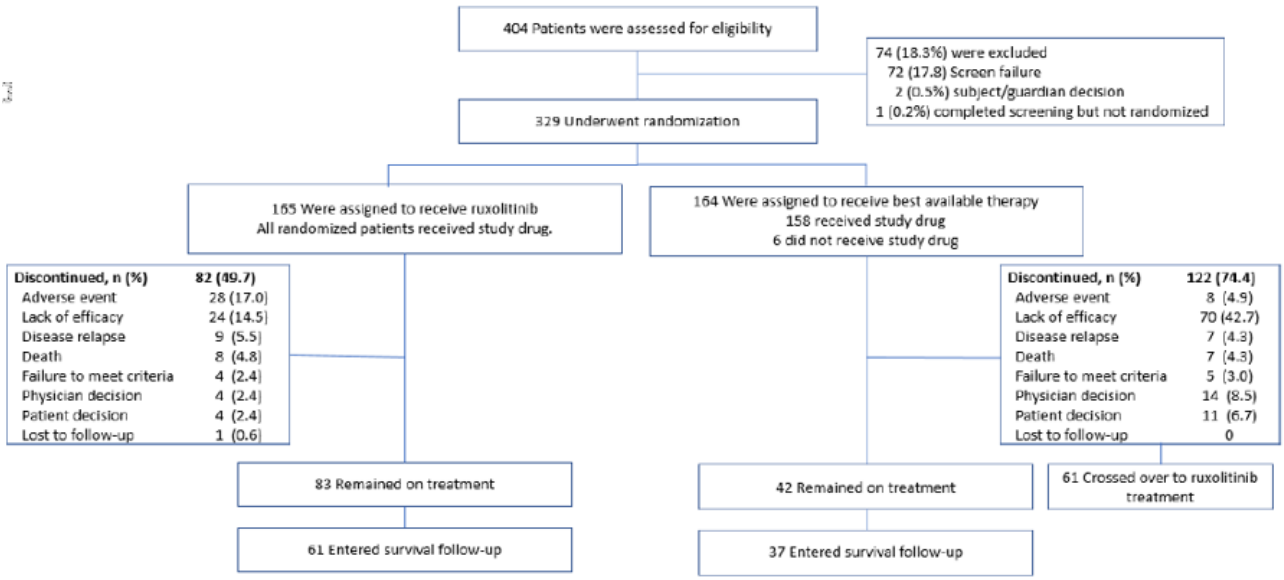
Screening failures

Table 14.1-1.3 (Page 1 of 1)
Screening phase disposition
All screened subjects

	All Subjects N=404 n (%)
Screened	404 (100)
Completed screening phase, randomized	329 (81.4)
Completed screening, not randomized	1 (0.2)
Did not complete screening	74 (18.3)
Primary reason for not completing screening	
Screen failure	72 (17.8)
Subject/guardian decision	2 (0.5)
Adverse event	0
Death	0
Disease relapse	0
Failure to meet protocol continuation criteria	0
Graft loss	0
Lack of efficacy	0
Lost to follow-up	0
Physician decision	0
Pregnancy	0
Protocol deviation	0
Study terminated by sponsor	0
Technical problems	0

Disposition of patients

Participant flow



Patient disposition - End of randomized treatment (FAS)

	RUX N=165 n (%)	BAT N=164 n (%)	All Patients N=329 n (%)
Subjects randomized			
Treated	165 (100)	158 (96.3)	323 (98.2)
Not treated	0	6 (3.7)	6 (1.8)
Treatment ongoing *	83 (50.3)	42 (25.6)	125 (38.0)
Discontinued from treatment period	82 (49.7)	122 (74.4)	204 (62.0)
Primary reason for discontinuation			
Adverse event	28 (17.0)	8 (4.9)	36 (10.9)
Lack of efficacy	24 (14.5)	70 (42.7)	94 (28.6)
Disease relapse	9 (5.5)	7 (4.3)	16 (4.9)
Death	8 (4.8)	7 (4.3)	15 (4.6)
Failure to meet protocol continuation criteria	4 (2.4)	5 (3.0)	9 (2.7)
Physician decision	4 (2.4)	14 (8.5)	18 (5.5)
Subject/guardian decision	4 (2.4)	11 (6.7)	15 (4.6)
Lost to follow-up	1 (0.6)	0	1 (0.3)
Continued into next phase at the end of randomized treatment			
Crossover treatment	0	61 (37.2)	61 (18.5)
Entered survival follow-up	61 (37.0)	37 (22.6)	98 (29.8)

*Ongoing treatment and/or assessments during the Main treatment period at cut-off date of 08-May-2020
Reason of discontinuation was also reported for subjects not treated

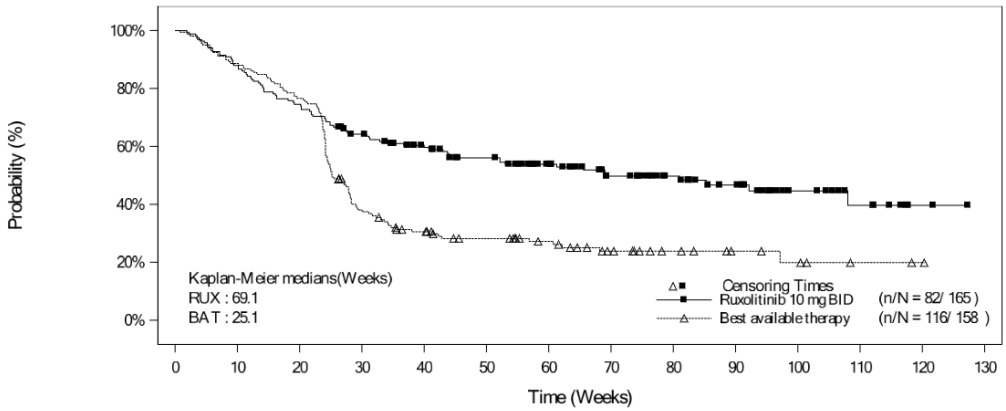
The median duration of exposure to RUX treatment and BAT treatment were 41.3 weeks (0.7-127.3) and 24.1 weeks (0.6-108.4) for RUX treatment and BAT treatment, respectively.

As illustrated in the K-M plot, treatments were discontinued at similar pace up to the measurement timepoint for the primary endpoint (Cycle 7 Day 1), after which there is a sharp drop in the BAT arm, explained by the study design as it was targeted to keep patients in the main treatment period at least until the Cycle 7 Day 1 visit.

A total of 82 (49.7%) and 122 (74.4%) patients discontinued the randomized treatment period in the ruxolitinib and BAT arm, respectively. The most frequent reasons for discontinuations during period up

to the Cycle 7 Day 1 visit were adverse events and lack of efficacy in ruxolitinib arm (21 and 13 patients, respectively), and lack of efficacy in the BAT arm (29 patients).

Figure HA2 14.1-1.5 (Page 1 of 1)
Kaplan-Meier estimate of time to discontinuation from main study period
Safety set



No. of patients still at risk

Time(Weeks)	0	10	20	30	40	50	60	70	80	90	100	110	120	130
Ruxolitinib 10 mg BID	165	145	123	102	87	75	61	45	35	26	14	8	2	0
Best available therapy	158	140	121	57	42	33	26	17	11	7	5	2	1	0

CINC424D2301- Primary CSR

Table HA2 14.1-1.4
Discontinuation reasons by time interval for the main study period, by treatment
Full analysis set

Treatment: Ruxolitinib 10 mg BID

	Weeks since treatment start							
Reason	**>0 - 8**	**>8 - 16**	**>16 - 24**	**>24 - 32**	**>32 - 40**	**>40 - 48**	**>48**	**any**
Primary reason for discontinuation								
Adverse event	8	8	5	3	1	1	2	28 (17.0%)
Death	2	3	1	0	1	1	0	8 (4.8%)
Disease relapse	1	1	1	2	1	0	3	9 (5.5%)
Failure to meet protocol continuation criteria	1	0	0	0	0	2	1	4 (2.4%)
Lack of efficacy	1	7	5	6	1	1	3	24 (14.5%)
Lost to follow-up	0	0	0	0	0	0	1	1 (0.6%)
Physician decision	1	0	1	1	0	0	1	4 (2.4%)
Subject/guardian decision	0	3	1	0	0	0	0	4 (2.4%)

CINC424D2301- Primary CSR

Table HA2 14.1-1.4
Discontinuation reasons by time interval for the main study period, by treatment
Full analysis set

Treatment: Best available therapy

	Weeks since treatment start							
Reason	**>0 - 8**	**>8 - 16**	**>16 - 24**	**>24 - 32**	**>32 - 40**	**>40 - 48**	**>48**	**any**
Primary reason for discontinuation								
Adverse event	2	0	3	2	0	0	1	8 (4.9%)
Death	4	0	1	2	0	0	0	7 (4.3%)
Disease relapse	1	2	1	3	0	0	0	7 (4.3%)
Failure to meet protocol continuation criteria	1	1	1	0	1	0	0	5 (3.0%)
Lack of efficacy	3	8	18	29	7	3	2	70 (42.7%)
Lost to follow-up	0	0	0	0	0	0	0	0
Physician decision	0	2	3	5	0	0	1	14 (8.5%)
Subject/guardian decision	3	2	3	0	0	0	1	11 (6.7%)

Cross-over patients

At the end of Cycle 6, 61 (37.2%) patients in the BAT arm crossed over to RUX treatment and 37 (22.6%) entered the Survival follow-up phase. No patient completed the Cross-over treatment period with RUX at the time of data cut-off, and 46 (75.4%, out of 61 patients) were still receiving RUX. Therefore, efficacy data for the cross-over treatment population are not presented in the present CSR.

Fifteen (24.6%) patients discontinued the Cross-over treatment period with RUX due to AEs: lack of efficacy each in 4 (6.6%) patients, physician decision and patient/guardian decision each in 3 (4.9%) patients and death in one patient.

Conduct of the study

Protocol changes

The original study protocol dated 14 Mar 2017 was amended once on 20-Dec-2017. The main purpose of the amendment was to extend the available patient population. Due to protracted enrollment and feedback from health authorities and investigators, several inclusion and exclusion criteria were revised:

- The window for patients currently receiving systemic or topical corticosteroids for the treatment of cGvHD was extended from six months to twelve months prior to Cycle 1 Day 1.
- Instead of excluding patients that received any systemic treatment for cGvHD in addition to corticosteroids ± CNI, patients who have received one prior systemic treatment in addition to corticosteroids ± CNI were eligible as well.
- The criterion regarding overlap syndrome was clarified to exclude those patients with active acute progressive disease.
- The definition of severely impaired renal function was updated by adding the text "having estimated creatinine clearance <30 ml/min measured or calculated by Cockroft Gault equation (confirmed within 48 hours prior to study treatment start)". This criterion was updated to provide a more sensitive measure of mild or moderate kidney injury at study entry.
- Updated cholestatic criterion to add total bilirubin >2mg/dL attributable to GvHD. This criterion was added as a more specific exclusion of patients with ongoing hepatic acute-like GvHD
- Added impairment of GI function example of diarrhea attributed to GvHD. This criterion was added as a more specific exclusion of patients with ongoing acute-like GI GvHD.

Additionally, with the recent approval of ibrutinib by the FDA the original protocol-defined list of BAT options would not allow for a complete comparative assessment versus RUX. Therefore, ibrutinib was added to this list to reflect the complete list of treatment options for this patient population.

The Patient Global Impression of Severity (PGIS) and the Patient Global Impression of Change (PGIC) patient-reported outcomes have been added as exploratory endpoint.

Lastly, a Data Monitoring Committee was added to include an independent review group.

Protocol deviations

Protocol deviations were observed in 63% of patients in the RUX arm and 66.5% in BAT arm. The most common deviation category was 'other deviations'(32.7% vs. 37.2%), primarily consisting of Lee Symptom Scale not collected (13.9% vs. 17.1%) and HBV and/or HCV viral load testing value missing at Cycle 1 Day 1 (9.7% vs. 7.9%).

Protocol deviations due to COVID-19 pandemic

The study was fully recruited before the out brake of the COVID-19 pandemic and all, except 6, patients had completed Cycle 7 Day 1 visit (primary endpoint). The conduct of the clinical trial during the COVID-19 pandemic followed the guidance of FDA and EMA. One of the changes was allowing

assessments to be done remotely (video, telephone) or at non-conventional location such as outside of site, when no other option was available.

Deviations were noted with respect to drug supply (5.5%), visits conducted outside of study site (8,2%), cGVHD staging assessment missing (2,7%), assessment / procedure changed (2,7%), missing cGvHD assessment for cycle 6 and 7, resulting in "PD" (1.8%, 2 and 4 patients in the RUX arm and BAT arm, respectively). Due to challenges to assess GvHD patients, especially with regard to the staging of skin and lung involvement in cGvHD for patients with remote visits (video or telephone) during the COVID-19 pandemic, the applicant has declared that the majority of patients had already reached the time for primary endpoint assessment before the WHO declared COVID-19 as a pandemic. For the 8 patients, for whom the Cycle 7 Day 1 visit could have been impacted due to the pandemic, measures were taken to ensure a reliable assessment. Other visits done outside of study site due to COVID-19 were handled the same way.

Baseline data

Demographics and baseline characteristics (FAS)

Demographic variable	RUX N=165	BAT N=164	All Patients N=329
Age (years)			
n	165	164	329
Mean (SD)	45.9 (15.68)	47.2 (16.17)	46.5 (15.92)
Median	49.0	50.0	49.0
Q1-Q3	33.0-57.0	34.5-61.0	34.0-60.0
Min - Max	13.0-73.0	12.0-76.0	12.0-76.0
Age category -n (%)			
Adolescents, 12 - <18 years	4 (2.4)	8 (4.9)	12 (3.6)
18 - 65 years	143 (86.7)	134 (81.7)	277 (84.2)
>65 years	18 (10.9)	22 (13.4)	40 (12.2)
Sex -n (%)			
Female	56 (33.9)	72 (43.9)	128 (38.9)
Male	109 (66.1)	92 (56.1)	201 (61.1)
Race -n (%)			
White	116 (70.3)	132 (80.5)	248 (75.4)
Black or African American	2 (1.2)	0	2 (0.6)
Asian	33 (20.0)	21 (12.8)	54 (16.4)
American Indian or Alaska Native	2 (1.2)	0	2 (0.6)
Other	9 (5.5)	4 (2.4)	13 (4.0)
Unknown	3 (1.8)	7 (4.3)	10 (3.0)
Ethnicity -n (%)			
Hispanic/Latino	13 (7.9)	13 (7.9)	26 (7.9)
Not Hispanic/Latino	118 (71.5)	115 (70.1)	233 (70.8)
Not Reported	26 (15.8)	25 (15.2)	51 (15.5)
Unknown	8 (4.8)	11 (6.7)	19 (5.8)
Weight (kg)			
n	165	163	328
Mean (SD)	68.5 (18.29)	67.9 (16.71)	68.2 (17.50)
Median	66.0	67.0	66.1
Q1-Q3	55.0-78.9	55.0-78.5	55.0-78.6
Min - Max	32.0-128.0	37.0-128.5	32.0-128.5
Height (cm)			
n	143	150	293
Mean (SD)	169.7 (9.77)	169.4 (10.05)	169.6 (9.90)
Median	169.0	170.0	170.0
Q1-Q3	162.0-178.0	160.0-177.0	162.0-177.5
Min - Max	145.0-191.0	144.3-196.0	144.3-196.0
Body mass index (kg/m2)			
n	143	150	293
Mean (SD)	23.4 (5.35)	23.5 (4.92)	23.4 (5.13)

Demographic variable	RUX N=165	BAT N=164	All Patients N=329
Median	22.5	22.8	22.7
Q1-Q3	19.2-26.3	19.8-26.2	19.6-26.2
Min - Max	13.0-38.7	14.7-42.9	13.0-42.9
Assessment of performance status -n (%)			
ECOG	153 (92.7)	148 (90.2)	301 (91.5)
Karnofsky	41 (24.8)	43 (26.2)	84 (25.5)
Lansky	11 (6.7)	14 (8.5)	25 (7.6)
Missing	0	1 (0.6)	1 (0.3)
ECOG performance status -n (%)			
0	39 (23.6)	42 (25.6)	81 (24.6)
1	92 (55.8)	82 (50.0)	174 (52.9)
2	22 (13.3)	22 (13.4)	44 (13.4)
3	0	2 (1.2)	2 (0.6)
Missing	12 (7.3)	16 (9.8)	28 (8.5)
Karnofsky performance status -n (%)			
≥ 90	24 (14.5)	17 (10.4)	41 (12.5)
70 - 80	16 (9.7)	19 (11.6)	35 (10.6)
50 - 60	1 (0.6)	7 (4.3)	8 (2.4)
Missing	124 (75.2)	121 (73.8)	245 (74.5)
Lansky performance status -n (%)			
≥ 90	9 (5.5)	4 (2.4)	13 (4.0)
70 - 80	2 (1.2)	9 (5.5)	11 (3.3)
50 - 60	0	1 (0.6)	1 (0.3)
Missing	154 (93.3)	150 (91.5)	304 (92.4)

Body Mass Index (kg/m^2) = weight (kg) / (height (m))^2

Underlying disease history/transplant and GvHD related history

The most frequently reported underlying diseases were AML (36.5%), ALL (15.8%) and MDC 13.4%. The fraction of patients with a non-malignant underlying disease was low.

Using the CIBMTR risk assessment tool (a validated tool to categorize groups of patients undergoing allogeneic stem cell transplantation, intended for research purposes, to stratify patients in broad disease risk categories for retrospective or prospective studies) approximately 25% were graded in each of the Low, Intermediate and High-risk groups. ~21% was missing or unknown.

Median time from transplant to cGvHD diagnosis was 235.00 days (range: 20.0 – 8047.0) overall and similar between two arms.

The median time since diagnosis to randomization was 2.33 years.

The stem cell source was bone marrow in 13.3% and 18.9% of the patients in the RUX arm and the BAT arm, respectively and peripheral blood in 85.5% and 79.9% of the patients in the RUX arm and the BAT arm, respectively. Only two patients in each arm received cord blood. Most patients received grafts from identical HLA-matched donors (57.5%; 192/329). T-cell depletion was performed in 38 (11.4%) patients overall.

Graft versus Host disease history (Full analysis set)

Disease history	RUX N=165	BAT N=164	All Patients N=329
Prior aGvHD-n (%)			
Any	92 (55.8)	88 (53.7)	180 (54.7)
Grade I	25 (15.2)	30 (18.3)	55 (16.7)
Grade II	53 (32.1)	43 (26.2)	96 (29.2)
Grade III	14 (8.5)	12 (7.3)	26 (7.9)
Grade IV	0	3 (1.8)	3 (0.9)
Steroid-refractory aGvHD	18 (10.9)	17 (10.4)	35 (10.6)
Time from aGvHD diagnosis to resolution (days)			
n	90	83	173
Mean (SD)	143.12 (241.795)	105.78 (173.747)	125.21 (212.118)
Median	63.50	50.00	52.00
Min-Max	4.0 - 1675.0	5.0 - 1227.0	4.0 - 1675.0
Time from aGvHD diagnosis to randomization (days)			
n	92	87	179
Mean (SD)	578.76 (485.490)	631.05 (1110.281)	604.17 (846.623)
Median	454.00	370.00	416.00
Min-Max	110.0 - 2558.0	57.0 - 9981.0	57.0 - 9981.0
Overall severity of cGvHD at initial diagnosis			
Mild	33 (20.0)	41 (25.0)	74 (22.5)
Moderate	77 (46.7)	77 (47.0)	154 (46.8)
Severe	53 (32.1)	45 (27.4)	98 (29.8)
Unknown	1 (0.6)	0	1 (0.3)
Missing	1 (0.6)	1 (0.6)	2 (0.6)
Time from transplant to cGvHD diagnosis (days)			
n	165	164	329
Mean (SD)	371.44 (378.120)	404.53 (749.580)	387.94 (592.439)

Disease history	RUX N=165	BAT N=164	All Patients N=329
Median	247.00	230.00	235.00
Min-Max	20.0 - 2360.0	35.0 - 8047.0	20.0 - 8047.0
Time from cGvHD diagnosis to randomization (days)			
n	165	164	329
Mean (SD)	232.62 (282.843)	227.24 (287.471)	229.94 (284.737)
Median	174.00	149.50	154.00
Min-Max	7.0 - 2017.0	10.0 - 1947.0	7.0 - 2017.0
SR-cGvHD diagnosis-n (%)			
SR criteria met (any)*	165 (100)	164 (100)	329 (100)
A: lack of response or disease progression after prednisone ≥ 1 mg/kg/day for at least 1 week (or equivalent)	62 (37.6)	73 (44.5)	135 (41.0)
B: Disease persistence without improvement despite continued treatment with prednisone >0.5 mg/kg/day or 1 mg/kg/every other day for at least 4 weeks (or equivalent)	58 (35.2)	42 (25.6)	100 (30.4)
C: Increase prednisone dose to >0.25 mg/kg/day after two unsuccessful attempts to taper the dose (or equivalent)	45 (27.3)	49 (29.9)	94 (28.6)
Time from initial cGvHD to diagnosis of SR cGvHD (days)			
n	165	164	329
Mean (SD)	200.84 (259.325)	186.87 (242.706)	193.88 (250.892)
Median	125.00	106.00	111.00
Min-Max	3.0 - 2009.0	2.0 - 1540.0	2.0 - 2009.0
Overall severity of SR-cGvHD at study entry			
Mild	0	1 (0.6)	1 (0.3)
Moderate	68 (41.2)	73 (44.5)	141 (42.9)
Severe	97 (58.8)	90 (54.9)	187 (56.8)
Prior systemic cGvHD / SR-cGvHD therapy-n (%)			
Steroid only	70 (42.4)	81 (49.4)	151 (45.9)
Steroid + CNI	68 (41.2)	69 (42.1)	137 (41.6)
Steroid + CNI + other systemic therapy	10 (6.1)	4 (2.4)	14 (4.3)
Steroid + other systemic therapy	14 (8.5)	9 (5.5)	23 (7.0)
Missing	3 (1.8)	1 (0.6)	4 (1.2)

*: Data were reported by investigator.
-Analyses of 'time to variables' are given for patients with available start and end dates only
-Prior treatment for cGvHD as documented in the prior medication data, topical or local treatments not counted

Table 14.1-7.5 (Page 1 of 1)
Chronic GvHD Staging at screening
Full analysis set

	RUX N=165	BAT N=164	All Patients N=329
Overall cGvHD Severity -n (%) [1]			
Mild	1 (0.6)	1 (0.6)	2 (0.6)
Moderate	67 (40.6)	74 (45.1)	141 (42.9)
Severe	97 (58.8)	89 (54.3)	186 (56.5)
Organ involvement -n (%) [2]			
Mouth	97 (58.8)	103 (62.8)	200 (60.8)
Gastrointestinal	39 (23.6)	36 (22.0)	75 (22.8)
Lung	74 (44.8)	67 (40.9)	141 (42.9)
Eyes	97 (58.8)	92 (56.1)	189 (57.4)
Joints And Fascia	48 (29.1)	42 (25.6)	90 (27.4)
Liver	42 (25.5)	40 (24.4)	82 (24.9)
Skin	121 (73.3)	113 (68.9)	234 (71.1)
Genital Tract	14 (8.5)	17 (10.4)	31 (9.4)
Missing	0	1 (0.6)	1 (0.3)

With regard to prior aGVHD diagnosis, this was documented for in total 189 (54.7%) patient with 38% with Grade ≥2. Out of these, 10.6% had a history of SR-aGvHD. The median time form aGvHD diagnosis and resolution was 52 days and the median time form resolution of aGvHD to randomization was 416 d.

Medical history and cGvHD disease history

Most reported ongoing medical conditions were consistent with the study indication and the complications of treatment (transplantation and GvHD treatment).

Prior prophylaxis - Among the 329 patients randomised, 33.9% of patients in RUX arm and 31.1% of patients in BAT arm received prior prophylaxis treatment for cGvHD. CNIs were the most frequently (24.0%) reported prophylactic therapy for cGvHD in both arms (23.6% and 24.4%, respectively), including cyclosporine (14.5% and 17.1%) and tacrolimus (9.1% and 7.9%). Use of glucocorticoids in the prophylactic treatment for cGvHD was reported in 10.0% of all patients.

Prior treatment - Almost all patients reported a prior treatment for cGvHD or SR-cGvHD (RUX arm: 98.2% and BAT arm: 95.7%). The most frequently used prior systemic therapy was steroid alone (42.4% in RUX arm and 49.4% in BAT arm) or steroids + CNI (41.2% vs. 42.1%).

Concomitant cGvHD treatments – Almost all (≥99%) patients in both treatment arms used concomitant immunosuppressive medications on treatment (starting on or after the start of study treatment and no more than 30 days after EOT). Corticosteroids were used by 59.4% of patients in the RUX arm and 55.1% in the BAT arm. Calcineurin inhibitors were taken in 27.3% vs. 19%, respectively.

Transfusions

A total of 26.7% patients in the ruxolitinib arm and 21.5% patients in the BAT arm underwent transfusion starting on or after the start of study treatment (packed red blood cells in 23.6% and 18.4% in each respective arm) and platelets (13.9% in each arm).

Numbers analysed

Analysis sets (all randomized patients)

Efficacy analyses used the Full Analysis Set (FAS), following the Intent-to-Treat (ITT) principle, comprising all patients to whom study treatment has been assigned by randomisation. All 329 randomised patients were included in the Full Analysis Set (FAS) for the primary analysis with data cut-off date 08-May-2020: n=165 in the RUX arm and n=164 in the BAT arm.

Supportive analyses were performed using the Per Protocol Set comprised of patients compliant with the requirements of the study (n=96 [58.2%] in the RUX arm, and n=92 [56.1%] in the BAT arm).

Results of the interim analysis (IA) for the primary and key secondary efficacy endpoints were based on the first 196 patients who completed Cycle 7 Day 1 visit (by data cut-off on 09-Jul-2019), after which the DMC recommended to proceed the study with primary analysis for all 329 randomized patients who completed Cycle 7 Day 1 visit or discontinued earlier (data cut-off on 08-May-2020).

Outcomes and estimation

Statistical tests of the primary and the two key secondary endpoints at the interim analysis and the primary analysis were performed according to an overall hierarchical testing procedure (please refer to Statistical methods).

Primary efficacy results

Primary endpoint, overall response rate (ORR) at Cycle 7 Day 1, as assessed by local investigators' review of cGvHD response according to the NIH response criteria for IA (interim analysis) and primary analysis.

Overall response rate at Cycle 7 Day 1 (Full analysis set – Interim Analysis)

	RUX N = 97		BAT N = 99		Odds ratio	95% CI	P-value
	n (%)	95% CI	n (%)	95% CI			
Responders							
Complete Response (CR)	8 (8.2)		3 (3.0)				
Partial Response (PR)	41 (42.3)		23 (23.2)				
Non-responders							
Unchanged Response	5 (5.2)		8 (8.1)				
Mixed response	7 (7.2)		11 (11.1)				
Progression	2 (2.1)		15 (15.2)				
Other *	3 (3.1)		4 (4.0)				
Unknown	31 (32.0)		35 (35.4)				
Death	9 (9.3)		7 (7.1)				
Early discontinuation	18 (18.6)		21 (21.2)				
Missing visits	4 (4.1)		7 (7.1)				
Overall Response Rate (ORR: CR+PR)	49 (50.5)	(40.2, 60.8)	26 (26.3)	(17.9, 36.1)	2.98	(1.62, 5.48)	0.0003

N: The total number of subjects in the treatment group. It is the denominator for percentage (%) calculation.
n: Number of subjects who are at the corresponding category.
The 95% CI for the response rate was calculated using Clopper Pearson exact method.
One-sided p-value, odds ratio and 95% CI are calculated using stratified CMH test.
* Other: subjects with additional systemic therapies along with CR/PR per investigator assessment

The results were confirmed at the primary analysis (data cut-off date 08-May-2020, N=329).

Overall response rate at Cycle 7 Day 1 (Full analysis set– Primary analysis)

	RUX N = 165		BAT N = 164		Odds ratio	95% CI	P-value
	n (%)	95% CI	n (%)	95% CI			
Overall response							
Responders							
Complete Response (CR)	11 (6.7)		5 (3.0)				
Partial Response (PR)	71 (43.0)		37 (22.6)				
Non-responders							
Unchanged Response	9 (5.5)		15 (9.1)				
Mixed response	10 (6.1)		17 (10.4)				
Progression	4 (2.4)		21 (12.8)				
Other *	5 (3.0)		9 (5.5)				
Unknown	55 (33.3)		60 (36.6)				
Death	16 (9.7)		11 (6.7)				
Early discontinuation	33 (20.0)		33 (20.1)				
Missing visits	6 (3.6)		16 (9.8)				
Overall Response Rate (ORR: CR+PR)	82 (49.7)	(41.8, 57.6)	42 (25.6)	(19.1, 33.0)	2.99	(1.86, 4.80)	<0.0001

N: The total number of subjects in the treatment group. It is the denominator for percentage (%) calculation.
n: Number of patients who are at the corresponding category.
The 95% CI for the response rate was calculated using Clopper Pearson exact method.
One-sided p-value, odds ratio and 95% CI are calculated using stratified CMH test.

It is noticed that ORR rates are considerably lower than what was anticipated in the sample size calculation.

The P-value of the stratified CMH test (P<0.0001) is given for descriptive purpose only, as significance was shown at the interim analysis.

The proportion of complete responders was low in both arms (6.7% vs. 3%).

Sensitivity analysis using PPS (N=188), ORR at Cycle 7 Day 1 evaluated with the same analysis conventions as for the primary efficacy analysis, was consistent with the primary analysis using FAS.

Overall response at Cycle 7 Day 1 – (Per Protocol Set, primary analysis)

	RUX N = 96		BAT N = 92				
	n (%)	95% CI	n (%)	95% CI	Odds ratio	95% CI	p-value
Overall response							
Responders							
Complete Response (CR)	6 (6.3)		2 (2.2)				
Partial Response (PR)	47 (49.0)		20 (21.7)				
Non-responders							
Unchanged Response	5 (5.2)		12 (13.0)				
Mixed response	5 (5.2)		10 (10.9)				
Progression	2 (2.1)		13 (14.1)				
Other *	3 (3.1)		5 (5.4)				
Unknown	28 (29.2)		30 (32.6)				
Death	9 (9.4)		8 (8.7)				
Early discontinuation	15 (15.6)		15 (16.3)				
Missing visits	4 (4.2)		7 (7.6)				
Overall Response Rate (ORR: CR+PR)	53 (55.2)	(44.7, 65.4)	22 (23.9)	(15.6, 33.9)	4.08	(2.16, 7.69)	<0.0001

N: The total number of subjects in the treatment group. It is the denominator for percentage (%) calculation.
n: Number of patients who are at the corresponding category.
The 95% CI for the response rate was calculated using Clopper Pearson exact method.
One-sided p-value, odds ratio and 95% CI are calculated using stratified CMH test.
* Other: patient with additional systemic therapies along with CR/PR per investigator assessment

According to the SAP, a patient was not considered a responder at Cycle 7 Day 1 if any of the following events occur in case of missing cGvHD assessment at Cycle 7 Day 1, or in case there was use of Additional systemic therapy for cGvHD. It has been clarified that missing data in the primary analysis are presented in category 'unknown', while the category 'missing visits' actually refers to invalid cGvHD assessments. There were 19 (11.5%) and 49 (29.9%) patients in the ruxolitinib and BAT arm, respectively, who started new or additional systemic cGvHD treatment up to the pre-specified time window used to define a valid Cycle 7 Day 1. Sensitivity analysis was requested to use the actual response instead of imputed non-response for patients who received a new systemic treatment for cGvHD prior to Cycle 7 Day 1 visit. The analysis was not provided as done in REACH-2 study; instead, only frequency of the overall response was presented per treatment group: 87 (52.7%) and 51 (31.1%) patients with overall response in the ruxolitinib and BAT arm, respectively. With non-responder imputation in the primary analysis, the corresponding numbers were 82 (49.7%) and 42 (25.6%) patients with overall response in the ruxolitinib and BAT arms, respectively, thus, potentially overestimating efficacy of ruxolitinib.

Initial BAT treatment, assigned before randomization

Table 14.3-1.3 (Page 1 of 1)
Initial BAT treatment and number of BATs started before Cycle 7 Day 1
Safety set

	BAT N=158
Initial BAT treatment -n (%)	
Extracorporeal photopheresis (ECP)	55 (34.8)
Mycophenolate mofetil (MMF)	35 (22.2)
Ibrutinib	27 (17.1)
Low-dose Methotrexate (MTX)	10 (6.3)
Imatinib	8 (5.1)
Sirolimus	7 (4.4)
Rituximab	6 (3.8)
Everolimus	5 (3.2)
Infliximab	5 (3.2)
Number of BAT treatments started before Cycle 7 Day 1 -n (%)	
1	127 (80.4)
2	28 (17.7)
>2	3 (1.9)

ORR at Cycle 7 Day 1 by initial BAT[2]

	Insufficient BAT compliance monitoring (worst case) per inspection report	N BAT (%)*	ORR BAT at Cycle 7 Day 1		N Rux	ORR RUX at Cycle 7 Day 1	
			ORR (%)**	95% CI		ORR (%)	95% CI
Total	66	164	42 (25.6)	19.1–33.0	165	82 (49.7)	41.8–57.6
Everolimus	4	5 (3.0)	1 (20.0)	0.5–71.6			
ECP	2	55 (33.5)	16 (29.1)	17.6–42.9			
Ibrutinib	13	27 (16.5)	6 (22.2)	8.6–42.3			
Imatinib	5	8 (4.9)	2 (25.0)	3.2–65.1			
Infliximab	0	5 (3.0)	1 (20.0)	0.5–71.6			
Low dose MTX	9	10 (6.1)	3 (30.0)	6.7–65.2			
MMF	26	35 (21.3)	10 (28.6)	14.6–46.3			
Rituximab	0	6 (3.7)	1 (16.7)	0.4–64.1			
Sirolimus	7	7 (4.3)	2 (28.6)	3.7–71.0			

ATG anti-thymocyte globulin, ECP: extracorporeal photopheresis, MMF mycophenolate mofetil MSC mesenchymal stromal cells, MTX methotrexate, RUX: ruxolitinib,* uses all patients randomised to BAT as reference (denominator).

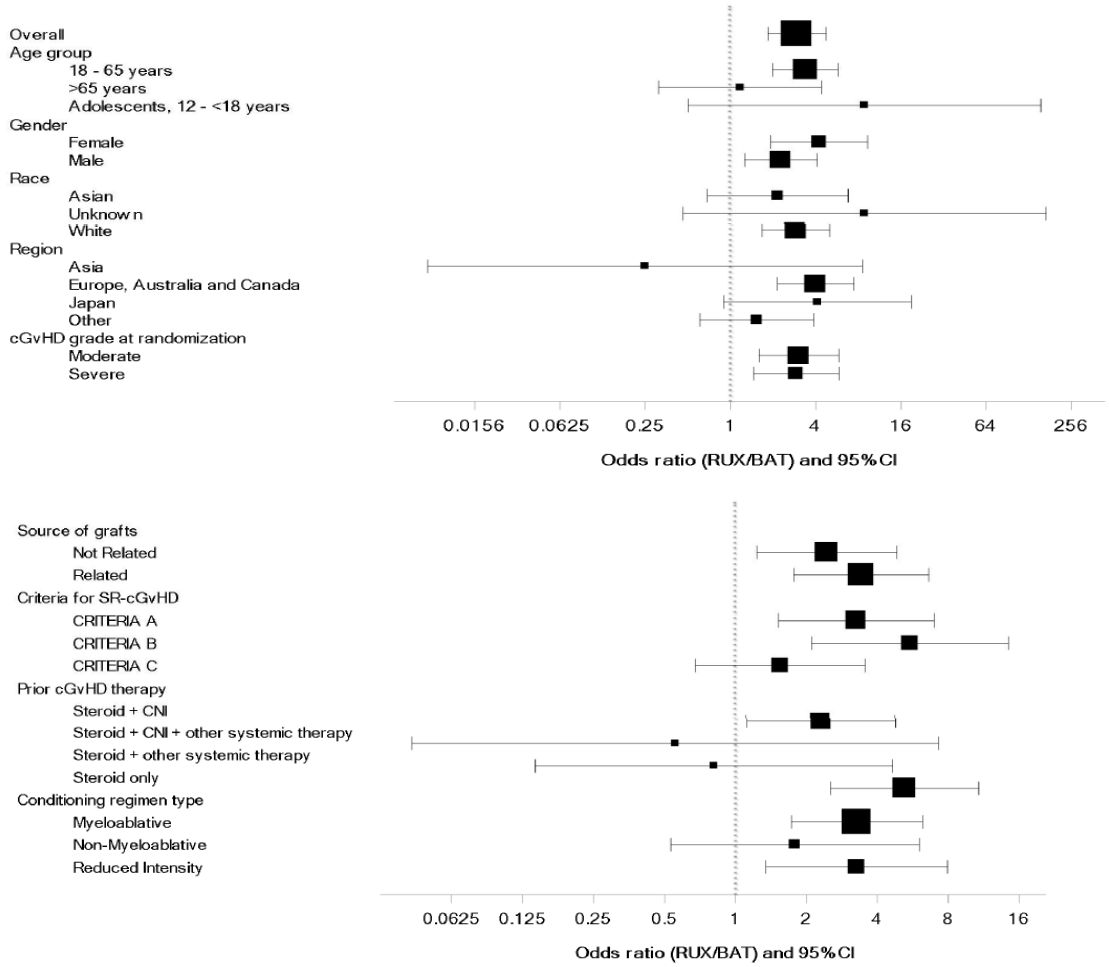
2 Supplement to: Zeiser R, Polverelli N, Ram R, et al. Ruxolitinib for glucocorticoid-refractory chronic graft-versus host disease. N Engl J Med 2021;385:228-38. DOI: 10.1056/NEJMoa2033122.

According to the SAP, a patient was not considered a responder at Cycle 7 Day 1 if any of the following events occurs in case of missing cGvHD assessment at Cycle 7 Day 1, or in case there was use of Additional systemic therapy for cGvHD. It has been clarified that missing data in the primary analysis are presented in category 'unknown', while the category 'missing visits' actually refers to invalid cGvHD assessments. There were 19 (11.5%) and 49 (29.9%) patients in the ruxolitinib and BAT arm, respectively, who started new or additional systemic cGvHD treatment up to the pre-specified time window used to define a valid Cycle 7 Day 1. Sensitivity analysis was requested to use the actual response instead of imputed non-response for patients who received a new systemic treatment for cGvHD prior to Cycle 7 Day 1 visit. The analysis was not provided as done in REACH-2 study; instead, only frequency of the overall response was presented per treatment group: 87 (52.7%) and 51 (31.1%) patients with overall response in the ruxolitinib and BAT arm, respectively. With non-responder imputation in the primary analysis, the corresponding numbers were 82 (49.7%) and 42 (25.6%) patients with overall response in the ruxolitinib and BAT arms, respectively, thus, potentially overestimating efficacy of ruxolitinib.

For the subgroup analyses, the odds ratio remained favorable for ruxolitinib, demonstrated by consistent odds ratio >1 in most subgroups. Subgroups with odds ratio <1 (Asian region, Steroid+ CNI+ other systemic therapy and steroid+ other systemic therapy), where all hampered by wide CI and small size. Point estimate of the treatment effect and 95% confidence intervals was provided.

In subgroup age 12- <18 years, there were in total 12 patients randomized, 4 patients in the RUX arm and 8 in the BAT treatment arm. Seventy five percent (3/4) and 25% (2/8) in the RUX arm and the BAT arm, respectively, showed a PR response. The odds ratio in subgroup adolescent patients were inconclusive due to the low number of evaluable patients.

Forest plot of odds ratio with 95% confidence interval for ORR at Cycle 7 Day 1 from subgroup analysis (Full analysis set – <u>Primary analysis</u>)



X-axis values are represented in natural log scale. Dotted lines shows no effect point.
The area of the box indicates the weight of the sub group, measured by the size of subpopulation.
Criteria for SR-cGvHD :
A. Lack of response or disease progression after administration of minimum prednisone 1 mg/kg/day for at least 1 week;
B. Disease persistence without improvement despite continued treatment with prednisone at >0.5 mg/kg/day or 1 mg/kg/every other day for at least 4 weeks;

C. Increase prednisone dose to >0.25 mg/kg/day after two unsuccessful attempts to taper the dose

Baseline organ involvement are overall balanced between the two treatment arms, with the exception of lung involvement (42.4% in the RUX arm and 29.9% in the BAT arm). Response assessments, i.e., responders at Cycle 7 Day 1, generally shows a higher response rate for all assessed organs in the RUX arm compared to the BAT arm.

Table 9. Individual Organ Response at Cycle 7 Day 1

	Ruxolitinib n=165		BAT n=164	
Organ	Baseline involvement[a]	Organ response[b]	Baseline involvement[a]	Organ response[b]
	n (%)	m/n (%)	n (%)	m/n (%)
Skin	119 (72.1)	49/119 (41.2)	110 (67.1)	17/110(15.5)
Eye	96 (58.2)	25/96 (26.0)	93 (56.7)	10/93 (10.8)
Mouth	96 (58.2)	48/96 (50.0)	99 (60.4)	25/99 (25.3)
Esophagus	18 (10.9)	9/18 (50.0)	17 (10.4)	5/17 (29.4)
Upper GI tract	20 (12.1)	8/20 (40.0)	21 (12.8)	8/21 (38.1)
Lower GI tract	15 (9.1)	8/15 (53.3)	10 (6.1)	3/10 (30.0)
Liver	86 (52.1)	21/86 (24.4)	83 (50.6)	18/83 (21.7)
Lung	70 (42.4)	6/70 (8.6)	49 (29.9)	3/49 (6.1)
Joints and fascia	45 (27.3)	17/45 (37.8)	44 (26.8)	7/44 (15.9)
Overall response	–	82 (49.7)	–	42 (25.6)

Organ response as documented by the investigator, m is the number of patients with organ response = CR or PR and excluding those responders where the organ abnormality is due to non-cGvHD reasons. Overall response counts patients with CR or PR as per investigator.
[a] Based on NIH cGvHD response guidelines (Lee SJ et al 2015)
[3] Baseline involvement if respective score at cycle 1 day 1 >0, or %FEV1 <75% (lung), ALT, bilirubin or AP > ULN (liver), joints and fascia score >0.
[b] m/n shows number of responders/patients with baseline involvement excluding in m those patients with change/addition of new systemic cGvHD treatment before Cycle 7 day 1

Cross-over population

At the primary analysis data cut-off (last patient enrolled could have reached the Cycle 7 Day 1 visit), 61 patients had crossed over from the BAT arm and started treatment with ruxolitinib, and treatment with ruxolitinib was still ongoing for 46 patients [Study D2301 PA Table 10-2]. Since not all of the 46 ongoing patients have reached the Cross-over Cycle 7 day 1 visit at the data cut-off, the analysis was performed again with longer follow-up based on the data cut-off 25-Jun-2021 used for the later OS analysis performed. In total 69 patients crossed-over from BAT to ruxolitinib; 33 of these patients had a response at cross-over Cycle 7 Day 1 corresponding to ORR of 47.8% (95% CI: 35.6%, 60.2%).

Overall, these results indicate a clinical benefit of ruxolitinib treatment also for SRcGvHD patients who failed treatment with BAT.

Secondary efficacy results

Failure free survival

The first key secondary endpoint, FFS, assessed the time from date of randomization to the earliest of i) relapse or recurrence of underlying disease or death due to underlying disease, ii.) non-relapse mortality, or iii.) addition or initiation of another systemic therapy for cGvHD

The study met the key secondary efficacy endpoint FFS at the interim analysis, with a HR of 0.315 (95% CI: 0.205, 0.486; one-sided p-value = <0.0001).

P-value derived from 1-sided stratified log-rank test, required significance level was alpha=0.01176. Hazard ratio obtained from stratified Cox model using cGvHD severity at randomization as strata.
NE: not estimated

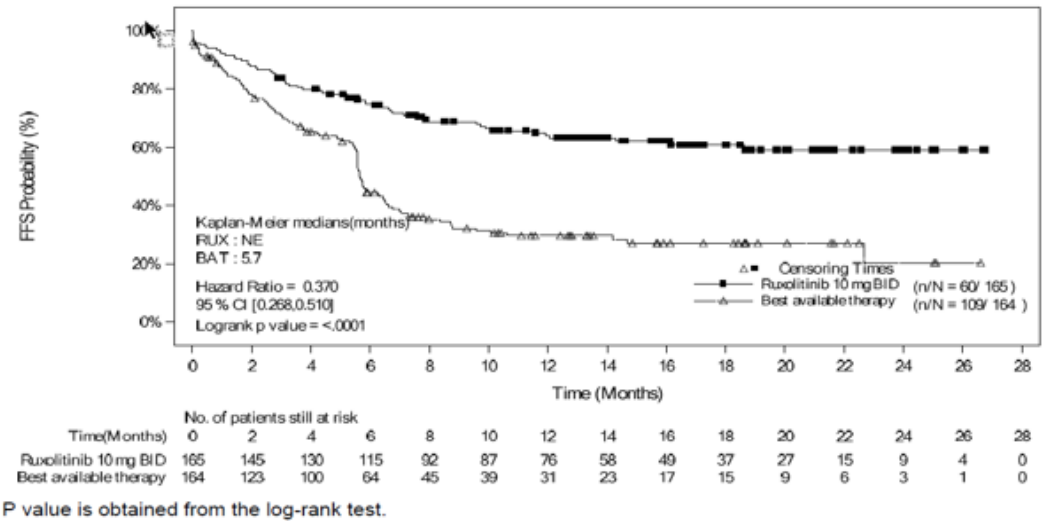
Kaplan-Meier estimate of failure free survival (Full analysis set -Primary analysis)

	RUX (N=165)	BAT (N=164)
Number of events	60 (36.4%)	109 (66.5%)
Number censored	105 (63.6%)	55 (33.5%)
Log-rank test P-value*	<0.0001	
Hazard ratio (95% CI)	0.370 (0.268, 0.510)	
Median FFS (months) (95% CI)	NE (18.6, NE)	5.7 (5.6, 6.5)
KM estimates (95% CI) at		
6 months	74.89 (67.48, 80.85)	44.46 (36.46, 52.14)
12 months	64.00 (55.78, 71.09)	29.62 (22.34, 37.23)
18 months	60.73 (52.06, 68.31)	27.04 (19.71, 34.88)
24 months	58.94 (49.80, 66.98)	20.28 (9.33, 34.19)

*: P-value derived from 1-sided stratified log-rank test, as statistical significance has met at interim analysis.
Hazard ratio obtained from stratified Cox model using cGvHD severity at randomization as strata

The superiority of FFS in the RUX arm compared to BAT was maintained at the primary analysis (n = 329), with HR=0.370 (95% CI: 0.268, 0.510; descriptive one-sided p<0.0001. The 6-month FFS probability was 74.89% (95% CI: 67.48, 80.85) in the RUX arm and was 44.46% (95% CI: 36.46, 52.14) in the BAT arm. The median FFS time was not reached for RUX and was 5.7 months for BAT.

Kaplan-Meier estimate of failure free survival (Full analysis set - Primary analysis)



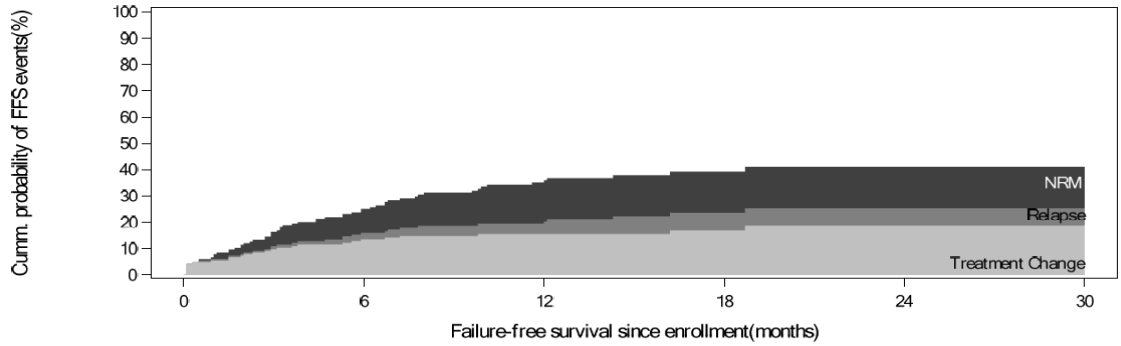
P value is obtained from the log-rank test.

A pronounced drop in the BAT curve is seen around month 6, supposedly triggered by cross over from BAT to RUX treatment. Patients in the BAT arm could cross over not earlier than Cycle 7 Day 1 visit and only if they were non-responders.

The FFS endpoint is a composite of (i) relapse or recurrence of underlying disease or death due of underlying disease, (ii) non-relapse mortality, or (iii) addition or initiation of another systemic therapy for cGvHD. As anticipated, this endpoint is driven by item (iii), while there is no indication of any impact of treatment on items (i) and (ii). The data are generated in an open-label study, where the impact of the cross-over option results in a precipitate fall in the KM curve for the control arm. However, this key secondary endpoint, shows statistical significance (HR of 0.315 (95% CI: 0.205, 0.486; one-sided p-value = <0.0001) in favor of ruxolitinib. The other items; NMR and MR, both each other's competing risk, had the similar cumulative incidence, hence gain in the one was not at the cost of the other.

Figure 14.2-1.2.1 Page (1 of 2)
Failure free survival by treatment
Full analysis set

Treatment=Ruxolitinib 10 mg BID

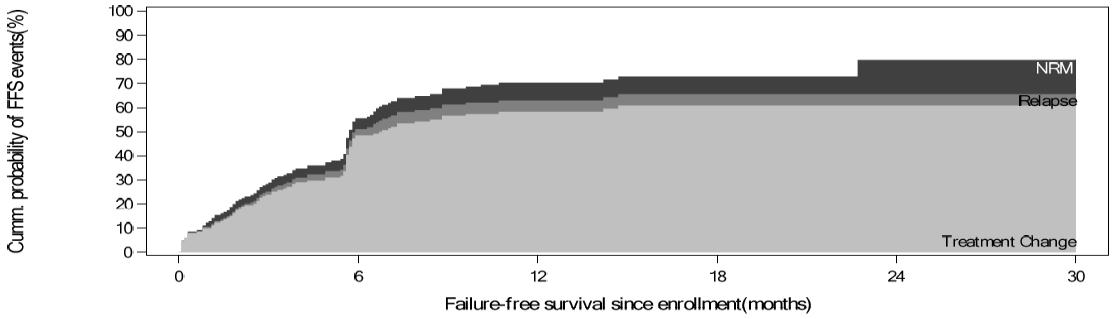


Treatment Change	0%	13.46%	15.55%	16.87%	18.66%	18.66%
Relapse	0%	2.46%	4.70%	6.65%	6.65%	6.65%
NRM	0%	9.19%	15.75%	15.75%	15.75%	15.75%
FFS	100%	74.89%	64.00%	60.73%	58.94%	58.94%

FFS probabilities were obtained by 100% - sum of the probabilities for Treatment change, Relapse and NRM.

Figure 14.2-1.2.1 Page (2 of 2)
Failure free survival by treatment
Full analysis set

Treatment=Best available therapy



Treatment Change	0%	48.51%	58.24%	60.81%	60.81%	60.81%
Relapse	0%	2.57%	4.69%	4.69%	4.69%	4.69%
NRM	0%	4.46%	7.46%	7.46%	14.22%	14.22%
FFS	100%	44.46%	29.61%	27.04%	20.28%	20.28%

Total symptom score, TSS; Patient reported outcomes based on the total symptom score

The second key secondary endpoint TSS, (rate of responders as per improvement ≥7 points of total TSS from baseline of the modified Lee Symptom Scale (mLSS), did not meet statistical significance (stratified CMH test p = 0.0151) during the interim analysis (based on N = 196). The results from the primary analysis are shown below.

Responders at Cycle 7 Day 1 based on the TSS by the mLSS (Full analysis set– Primary analysis)

	RUX N = 165		BAT N = 164				
	n (%)	95% CI	n (%)	95% CI	Odds ratio (RUX/BAT)	95% CI	p-value
Subjects with valid TSS at baseline	149 (90.3)		141 (86.0)				
Subjects with valid TSS at Cycle 7 Day 1							
All	92 (55.8)		87 (53.0)				
Without prior change of systemic cGvHD treatment	89 (53.9)		64 (39.0)				
Responders (TSS reduction >= 7 points)	40 (24.2)	(17.9, 31.5)	18 (11.0)	(6.6, 16.8)	2.62	(1.42, 4.82)	0.0011

Subjects with change of or addition of new systemic cGvHD treatment are counted as non-responders irrespective of the TSS value.

Significant improvement in TSS, (cycle 7 day 1) was shown for the RUX arm compared to the BAT arm at the primary analysis. The odds ratio (ruxolitinib/BAT) was 2.62 (95% CI: 1.42, 4.82, stratified CMH test, p = 0.0011).

Given the significant impact of symptoms on quality of life (QoL) in patients with cGvHD, the change in a modified Lee symptom score was assessed as one of the key secondary objectives. However, given the open label nature of the study, there is a risk of bias that cannot be ascertained as to its magnitude.

Therefore, TSS data are not included in the product information.

- *Other secondary endpoints*

Best overall response

BOR is defined as proportion of patients who achieved overall response (CR or PR) at any time point up to and including Cycle 7 Day 1 and before the start of additional systemic therapy for cGvHD. There was a statistically significant difference between the ruxolitinib and BAT (stratified CMH test p=0.0011, one-sided) with the odds ratio 2.17 (95% CI: 1.34, 3.52) for response in ruxolitinib arm compared to BAT arm. The rate of Non-responders in the RUX arm was 18.2% vs 28% in the BAT arm.

Best overall response up to Cycle 7 Day 1 (Full analysis set)

	RUX N = 165		BAT N = 164				
	n (%)	95% CI	n (%)	95% CI	Odds ratio	95% CI	p-value
Overall response							
Responders							
Complete Response (CR)	20 (12.1)		11 (6.7)				
Partial Response (PR)	106 (64.2)		88 (53.7)				
Non-responders							
Unchanged Response	27 (16.4)		33 (20.1)				
Mixed response	3 (1.8)		8 (4.9)				
Progression	0		5 (3.0)				
Unknown	9 (5.5)		19 (11.6)				
Death	1 (0.6)		1 (0.6)				
Early discontinuation	0		6 (3.7)				
Missing visits	8 (4.8)		12 (7.3)				
Overall Response Rate (ORR: CR+PR)	126 (76.4)	(69.1, 82.6)	99 (60.4)	(52.4, 67.9)	2.17	(1.34, 3.52)	0.0011

N: The total number of subjects in the treatment group. It is the denominator for percentage (%) calculation.
n: Number of patients who are at the corresponding category.
The 95% CI for the response rate was calculated using Clopper Pearson exact method.
One-sided p-value, odds ratio and 95% CI are calculated using stratified CMH test.
Unknown: no valid post-baseline response assessment

BOR up to Cycle 7 Day 1 (RUX arm 76.4% vs BAT arm 60.4%) compared to the primary endpoint ORR at Cycle 7 Day 1 (RUX arm 49.7% vs BAT arm 25.6%) show a greater rate of loss of responses in the BAT arm than the RUX arm.

While there is a high BOR rate (ORR [CR + PR] at any time point up to and including Cycle 7 Day 1) observed in both treatment arms, it is also noticed a relatively high rate of "lost responses" during the first 6 cycles (ruxolitinib and BAT treatment) as measured by the primary endpoint (i.e. ORR at Cycle 7, Day 1). However, early discontinuation, deaths, changes in systemic cGvHD treatment before Cycle 7 Day 1 or missing visits were the main reasons for the differences in BOR and ORR therefore, not necessarily meaning loss of response in all cases or treatment failure. For the RUX arm, 34.9% with BOR (CR or PR) had `no response´ at Cycle 7 D1 and the corresponding number for the BAT arm was 57.5%.

In addition, the probability of maintaining a response 12 months from the first documented response was 68.5% in the RUX arm and 40.3% in the BAT arm.

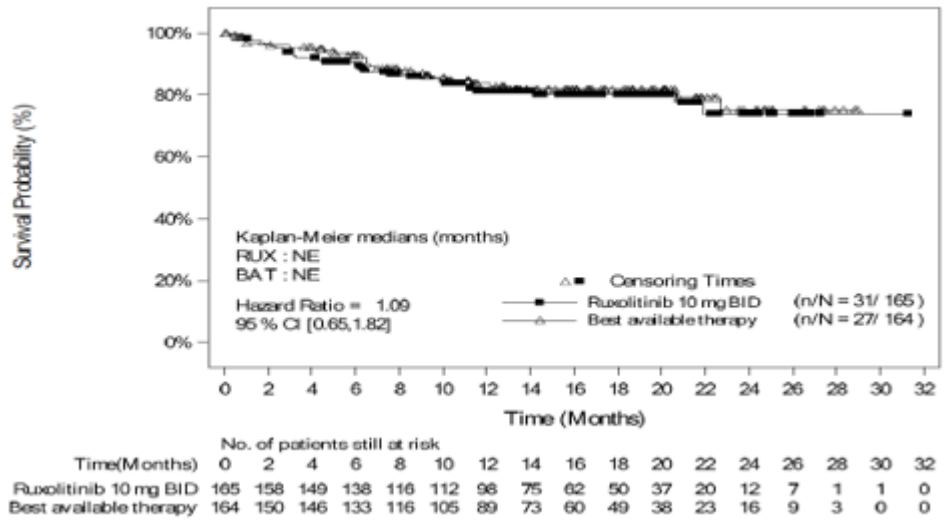
Best overall response during Cross-over treatment with ruxolitinib

In the Cross-over analysis set, 44 out of 61 patients had a duration of exposure to ruxolitinib longer than 24 weeks, i.e., completed 6 cycles of treatment. The ORR for the cross-over treatment period was 78.7% compared to 74.4% ORR for the RUX arm in the randomized treatment period.

Overall survival

At the time of primary analysis there were 31 deaths in the RUX arm and 27 deaths in the BAT arm (based on Cycle 7 Day 1 data). No statistical difference was observed between the arms (HR = 1.086 (95% CI: 0.648, 1.820), (log-rank p-value: 0.3764). The median OS estimated by Kaplan-Meier (K-M) was not reached at time of the analysis in either treatment arm.

Kaplan-Meier curves of overall survival (Full analysis set)



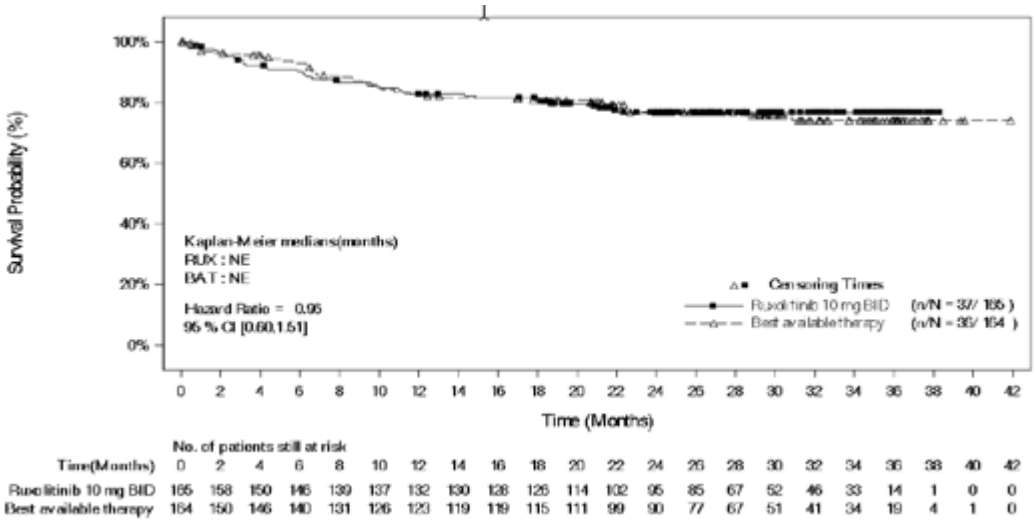
Given the study design, OS is primarily viewed as a safety endpoint. OS data are rather immature at the primary endpoint. An updated analysis of OS data has also been provided.

Overall survival results in REACH-3

Data cut-off	Events (Deaths) n (%)		Hazard ratio (Rux vs BAT)		Logrank test
	Rux N=165	BAT N=164	HR	95% CI	P-value*
08-May-2020	31 (18.8)	27 (16.5)	1.086	(0.648, 1.820)	0.3764
25-Jun-2021	37 (22.4)	36 (22.0)	0.956	(0.604, 1.512)	0.4230

*Nominal p-value derived from one-sided stratified logrank test with cGvHD severity as strata

Kaplan-Meier estimate of overall survival



The results remain immature, as is expected in the setting of cGvHD. Updated HR for OS shows a HR just below 1 (0.956 vs. 1.086 with previous analysis) with a CI still crossing 1. Updated results for primary cause of death did not show any new safety signals compared to the primary analysis results.

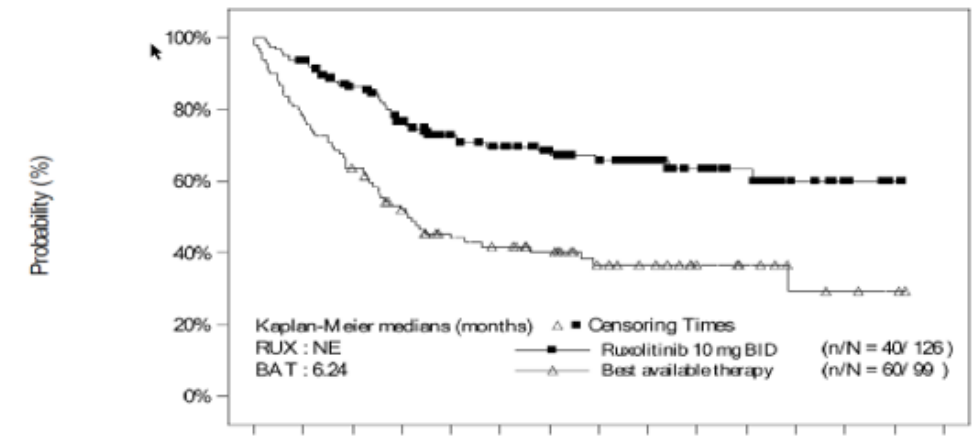
It is agreed that a detrimental effect on OS of ruxolitinib cannot be concluded based on available immature data. Mature OS data should be submitted post-marketing.

The allowance of cross-over creates a bias towards unity with regards to the impact of treatment on OS. In REACH-3 study, 61 (37.2%) of 164 subjects switched from BAT to ruxolitinib after completion of Cycle 6. The rank-preserving structural failure time (RPSFT) model was applied post-hoc to investigate what would have been the OS analysis had patients not crossed over to ruxolitinib. After adjustment for cross-over with the RPSFT model, the HR estimate was 0.967 (95% CI: 0.692, 1.352) compared to 0.956 (95% CI: 0.604, 1.512) obtained in the pre-planned ITT analysis. Thus, the RPSFT results do not indicate any impact of cross-over on overall survival.

Duration of Response

DOR was only measured in patients who achieved a CR or PR at or before Cycle 7 Day 1 (BOR). Median DOR was not reached in the RUX arm, while median DOR in the BAT arm was 6.2 months (95% CI: 4.7 to 13.3). Furthermore, the 6-, 12-, 18- and 24-months probability of maintaining a response, was in favor of the RUX arm.

Kaplan-Meier curve of duration of response (Full analysis set)



No. of patients still at risk

Time(Months)	0	2	4	6	8	10	12	14	16	18	20	22	24	26	28
Ruxolitinib 10 mg BID	126	117	101	85	71	63	53	46	34	24	18	9	5	1	0
Best available therapy	99	78	62	47	36	33	28	19	16	10	8	4	3	2	0

Include patients with BOR=CR and PR only

Table 14.2-1.8.1 (Page 1 of 1)
Duration of response
Full analysis set

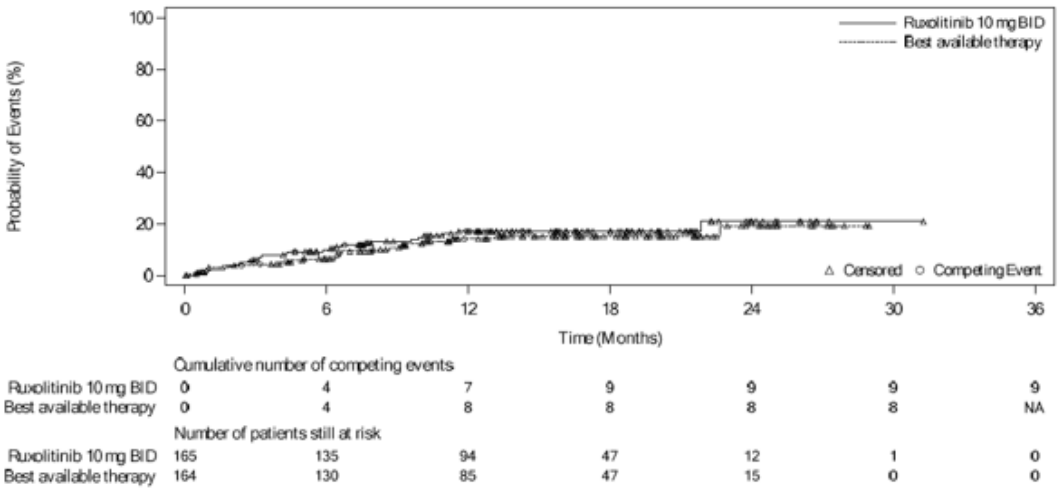
	RUX (N=126)	BAT (N=99)
Number of events	40 (31.7%)	60 (60.6%)
Number censored	86 (68.3%)	39 (39.4%)
Median DOR (months) (95% CI)	NE (20.2, NE)	6.2 (4.7, 13.3)
KM estimates (95% CI) at		
6 months	76.58(67.87, 83.22)	52.11(41.78, 61.45)
12 months	68.48(58.94, 76.26)	40.33(30.28, 50.15)
18 months	63.50(52.82, 72.38)	36.66(26.47, 46.88)
24 months	59.97(47.58, 70.33)	29.33(15.32, 44.84)

Non-relapse mortality

NRM curves for the RUX and BAT arms were overlapping (cumulative incidence) indicating similar event rates over time. The competing risk (hematological disease relapse/progression) was low in both treatment arms (9/165 and 8/164).

The number of patients censored was high in both treatment arms (129/165 and 134/164), indicating a high proportion of patients who were alive and had no relapse/progression.

Cumulative incidence of Non-Relapse Mortality (NRM) by treatment (Full analysis set)



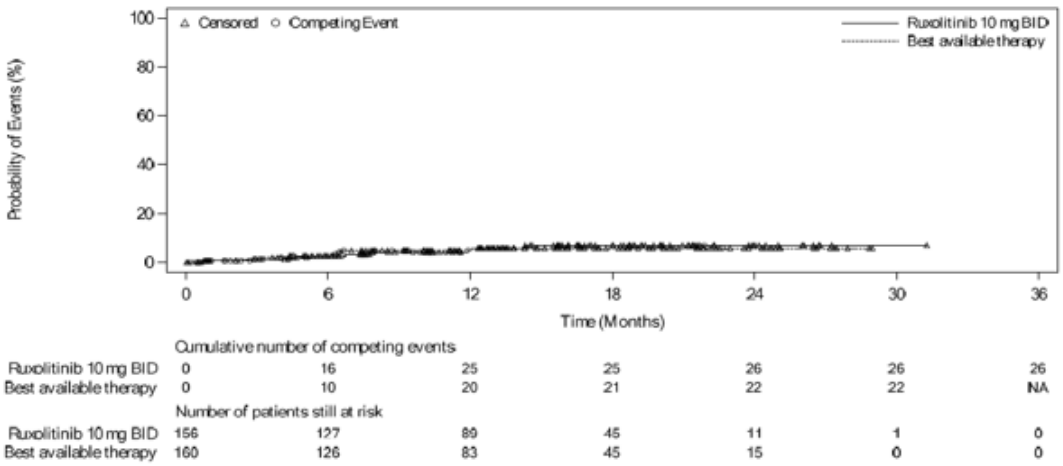
Cumulative number of competing events							
Ruxolitinib 10 mg BID	0	4	7	9	9	9	9
Best available therapy	0	4	8	8	8	8	NA
Number of patients still at risk							
Ruxolitinib 10 mg BID	165	135	94	47	12	1	0
Best available therapy	164	130	85	47	15	0	0

Incidence of Malignancy Relapse/ Recurrence

There were 156 patients in the ruxolitinib arm and 160 patients in the BAT arm who had underlying malignant disease at baseline. With regard to these patients, events of malignancy relapse/progression were low in both arms (9/156 in the ruxolitinib arm; 8/160 in the BAT arm).

Estimated cumulative event rates were 2.59% (95%CI: 0.85, 6.08) and 2.65% (95%CI: 0.87, 6.21) in respective treatment arms at 6 months and 4.94% (95%CI: 2.16, 9.45) and 5.80% (95%CI: 2.68, 10.65) at 12 months. Thus, indicating that the choice of cGvHD treatment has little or no significance for malignancy relapse, however data are still immature.

Cumulative incidence of Malignancy relapse/recurrence by treatment (Full analysis set)



Reduction of daily corticosteroid dose and successful tapering of all corticosteroids treatment

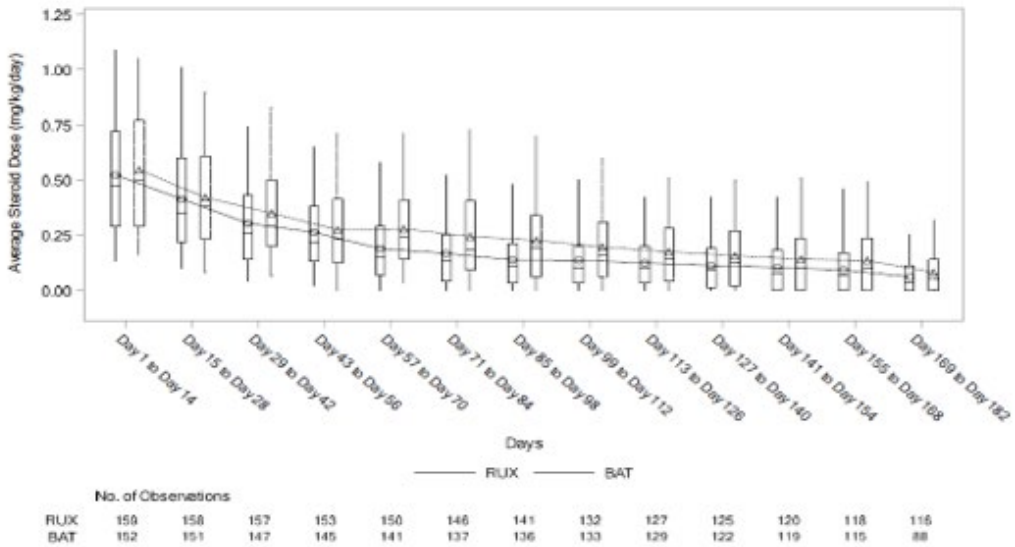
Tapering of corticosteroids was not recommended before 2 weeks after achieving a CR and tapering of CNI and/or ruxolitinib was not to be initiated until the patient was off corticosteroids and completed the assessments for Cycle 7 Day 1.

A comparable proportion of patients in both treatment arms had prior systemic therapy with steroids + CNI (RUX arm: 41.2%, BAT arm: 42.1%). As a consequence of successful tapering of steroids, subsequent tapering of CNI was to be initiated.

During the time interval of Day 155 to Day 168 (end of Cycle 6), patients with≥ 50% reduction of corticosteroid dose (body weight-normalized) from baseline occurred in 71.2% (84 out of 118) of patients in the ruxolitinib arm and in 69.6% (80 out of 115) of patients in the BAT arm. The mean dose intensity of steroid was lower in ruxolitinib than BAT, starting from Cycle 2 and the trend of slightly larger reduction of steroid dose in the ruxolitinib arm remained during the study.

Patients with no steroids during the interval occurred in 31.4% (37) of patients in the RUX arm and 27.8% (32) of patients in the BAT arm.

Average bi-weekly weight-standardized steroid up to Cycle 7 Day 1 (Safety set)



-Subjects who were completely tapered off steroids and were ongoing were counted as having steroid dose=0 until the end of the Main treatment period or the re-start of treatment with systemic steroids.
-Plot shows boxes (25th-75th percentiles) with median as horizontal line.
-The dots in the boxes and joint lines represent the mean values.
-Whiskers (vertical lines) extend to the 10th-90th percentiles.
-Values outside this range are not displayed.
-Dose of methylprednisolone was converted to prednisone equivalent

Data concerning duration of CNI treatment and change of CNI exposure by time interval up to Cycle 7 Day 1, by treatment arm and by type of CNI treatment (Cyclosporin or Tacrolimus) has been provided, showing similar exposure, between the arms, over time.

Patient-reported outcomes (FACT-BMT and EQ-5D-5L) and Patient reported outcomes – PGIS and PGIC

Although not defined as secondary endpoint, results of EQ-5D-5L and FACT-BMT questionnaires were presented. Completion rates as a percentage of available patients were similar between the two arms throughout the study.

Baseline scores were similar between both treatment arms. Higher scores of FACT-BMT assessment suggested a better the quality of life. The mean (SD) change from baseline at Cycle 7 Day 1 of the FACT-BMT score was 3.76 (15.028) in the RUX arm (n = 91) and 0.66 (16.816) in the BAT arm (n = 86). Change of EQ-5D-5L score from baseline was similar between the two arms, 0.07 (0.233) in the Rux arm (n = 90) vs. 0.00 (0.226) in the BAT arm (n = 84).

Exploratory efficacy endpoints: Patient reported outcomes – PGIS and PGIC

At Cycle 7 Day 1 visit, Patient global impression of severity (PGIS) values from patients that reported moderate, severe and very severe symptoms reduced as compared to Cycle 1 Day 1 in both arms. The PGIS values were lower in the RUX arm as compared to the BAT arm. The values for no and mild symptoms increased from Cycle 1 Day 1 visit.

Patient global impression of change (PGIC) values from patients, that reported "No change" at Cycle 2 Day 1, reduced at Cycle 7 Day 1 in the RUX arm from 20.9% to 7.1% and in the BAT arm from 28.6% to 22.4%. Numerically higher values of "moderately better" and "very much better" were reported in the RUX arm as compared to that in the BAT arm.

FACT-BMT up to Cycle 7, Day 1, measuring the mean change from baseline, a higher score suggests a better QoL. The mean change from baseline in the RUX arm is scores higher than that of the BAT arm.

Data from the mLSS scoring (please refer to key secondary endpoint) is by comparison of greater importance since this is a validated tool for cGvHD assessment.

The EQ-5D-5L mean score change from baseline up to Cycle 7 Day 1 was similar between both treatment arms.

There was a decrease in number of patients available for evaluation with time in both arms and more reduced in the BAT arm, due to a higher number of FFS event in the BAT arm and to cross over, which was allowed after Cycle 7 Day 1.

The The EQ-5D-5L and FACT-BMT were secondary endpoints and PGIC and PGIS were exploratory endpoints. The study was an open label study and none of these endpoints were controlled for multiplicity.

In summary, the PRO outcomes for the RUX arm compared to the BAT arm are difficult to evaluate.

Graft failure (Chimerism)

Only one patient, in the RUX arm, experienced graft failure and subsequently discontinued from study treatment, due to graft loss.

Summary of main studies

The following table summarise the efficacy results from the two main studies supporting the present application in aGvHD and cGvHD. These summaries should be read in conjunction with the discussion on clinical efficacy as well as the benefit risk assessment (see later sections).

Summary of Efficacy for trials CINC424C2301 (REACH2) and CINC424D2301 (REACH3)

Pivotal study 1: A phase III randomized open-label multi-center study of ruxolitinib versus best available therapy in patients with corticosteroid-refractory acute graft vs. host disease after allogeneic stem cell transplantation (INC424C2301)			
Study identifier	**REACH2**		
Design	Open label, randomized (1:1), phase III		
	Duration of main phase: Duration of Run-in phase: Duration of Extension phase:		24 weeks 0-28 days treatment taper could be delayed up to 2 years
Hypothesis	Superiority of Ruxolitinib over Best available therapy (BAT: anti-thymocyte globulin, extracorporeal photopheresis, mesenchymal stromal cells, low-dose methotrexate, mycophenolate mofetil, mTOR inhibitors [everolimus or sirolimus], etanercept or infliximab).		
Treatments groups	Experimental arm		Ruxolitinib 10 mg BID
	Control arm		BAT doses and administration were according to manufacturer's instructions and could be adjusted based on Investigator judgment
Endpoints and definitions	Primary endpoint	ORR day 28 (overall response rate)	proportion of patients in each arm demonstrating a complete response (CR) or partial response (PR) at Day 28 after randomization without requirement for additional systemic therapies for an earlier progression
	Key Secondary endpoint	Durable ORR day 56	proportion of all patients in each arm who achieved a complete response (CR) or partial response (PR) at Day 28 and maintain a CR or PR at Day 56.

Other sec endpoints are not corrected for multiplicity.	Other secondary endpoint	BOR (best overall response)	proportion of patients who had overall response, CR or PR, at any time point up to and including Day 28 and before the start of additional systemic therapy for aGvHD
	Other secondary endpoint	ORR day 14	
	Other secondary endpoint	DoR (duration of response)	evaluated in patients who achieved a CR or PR at or before Day 28
	Other secondary endpoint	Weekly cumulative steroid dose	cumulative steroid dose for each patient up to Day 56 or end of treatment
	Other secondary endpoint	OS (Overall survival)	the time from the date of randomization to the date of death due to any cause
	Other secondary endpoint	EFS (Event-free survival)	the time from the date of randomization to the date of hematologic disease relapse/progression, graft failure, or death due to any cause
	Other secondary endpoint	FFS (Failure-free survival)	hematologic disease relapse/progression, NRM or addition of new systemic aGvHD treatment, **and the competing risk was cGvHD**
	Other secondary endpoint	NRM (Non-relapse mortality)	the time from date of randomization to date of death not preceded by hematologic disease relapse/progression
	Other secondary endpoint	MR (Malignancy Relapse/Progression)	time from date of randomization to hematologic malignancy relapse/progression (only patients with underlying hematologic malignant disease)
	Other secondary endpoint	cGvHD	diagnosis of any cGvHD, including mild, moderate, severe
	Other secondary endpoint	OR up to and including Day 28	proportion of patients who achieved OR (CR+PR) at any time point up to and including Day 28 and before the start of additional systemic therapy for aGvHD
Database lock	DCO for primary analysis 25-Jul-2019 and for secondary analysis 06-Jan-2020		

Results and Analysis

Analysis description	**Primary Analysis**			
Analysis population and time point description	Full analysis set (=ITT)			
Effect estimate per comparison	Treatment group	**Ruxolotinib**	**BAT**	
	Number of subjects	154	155	
				Odds ratio (RUX/BAT)
	ORR d 28 (%)	62,3	39,4	2,64 (1.65, 4.22)
	(95% CI)	(54.2, 70,0)	(31.6, 47.5)	p<0.0001
	Durable OR Day 56 (%)	39.6	21.9	2,38 (1,43, 3,94)
	(95% CI)	(31.8, 47.8)	(15,7, 29,3)	p=0,0007

	ORR day 14	63.0	47.1	1,98 (1.24, 3.17)
	(95% CI)	(54.8, 70.6)	(39.0, 55.3)	p=0,0029*
	BOR up to d 28	81,8	60.6	3.07 1.80, 5.25)
	(95% CI)	(74.8, 87.6)	(52.5, 68.4)	p<0.0001*
Notes	*Results for ORR Day 14 and BOR are not corrected for multiplicity			
Analysis description	**Second analysis**			
Analysis population	Full analysis set (=ITT)			
Descriptive statistics/ estimate variability	Treatment group	**Ruxolotinib**	**BAT**	
	DoR (median/ Quartiles)	163 (78.0- 246.0)	101.0 (46.0-141.0)	Only subjects with CR/PR at d28
	OS (median/ months, 95% CI)	10.71	5.82	HR: 0.83 (CI 0.62, 1.13) p=0.2331*
	EFS (median/ months, 95% CI)	8.18	4.17	HR: 0.80 (CI 0.60, 1.08) p=0.1431*
	FFS (median/ months, 95% CI)	4.86	1.02	HR: 0.49, 95% CI: 0.37, 0.63; p<0.0001*
	MR (%)	10.9	17.0	
	NRM (%)	44.8	45.2	
	Completely tapered steroid dose by day 56 (%, CI 95%)	22 (15.8, 29.5)	14.8 (9.6, 21.4)	Odds Ratio: 1.63 (0.91, 2.92)
	Incid of cGvHD (%)	29.2	18.7	
Notes	*None of the secondary endpoints in the second analysis are corrected for multiplicity.			

<u>Pivotal study 2</u>: A phase III randomized open-label multi-center study of ruxolitinib versus best available therapy in patients with corticosteroid-refractory <u>chronic</u> graft vs. host disease after allogeneic stem cell transplantation.

Study identifier	**REACH3**	
Design	Open label, randomized (1:1) phase III	
	Duration of main phase: Duration of Run-in phase: Duration of Extension phase:	156 weeks (39 cycles) 0-28 days Cycle 7 to Cycle 39
Hypothesis	<u>Superiority of Ruxolitinib over Best available therapy</u> (BAT: Extracorporeal photopheresis, low-dose methotrexate, mycophenolate mofetil, mTOR inhibitors (everolimus or sirolimus), infliximab, rituximab, pentostatin, imatinib or ibrutinib	
Treatments groups	Experimental arm	Ruxolotinib 10 mg BID
	Control arm	BAT doses and administration were according to manufacturer's instructions and could be adjusted based on Investigator judgment

Endpoints and definitions	Primary endpoint	ORR (overall response rate)	the proportion of patients in each arm demonstrating a complete response (CR) or partial response (PR) without the requirement of additional systemic therapies for an earlier progression
	Co-Secondary endpoints	FFS (failure free survival)	time from date of randomization to the earliest of the following FFS events: 1) relapse or recurrence of underlying disease or death due to underlying disease or 2) non-relapse mortality or 3) addition or initiation of another systemic therapy for cGvHD.
		TSS (total symptom score)	the rate of patients with clinically relevant improvement of the modified Lee symptoms score at Cycle 7 Day 1 relative to baseline, as assessed by the rate of responders as per improvement ≥ 7 points of TSS from baseline of the modified Lee Symptom Scale
	Other secondary endpoint	BOR (Best overall response)	proportion of patients who achieved OR (CR+PR) at any time point (up Cycle 7 day 1 or the start of additional systemic therapy for cGvHD)
	Other secondary endpoint	OR at Cycle 4 Day 1	Proportion of patients who achieved OR (CR+PR) at Cycle 4 Day 1
	Other secondary endpoint	DOR (Duration of response)	for responders only; the time from first response until cGvHD progression, death, or the date of change/addition of systemic therapies for cGvHD
	Other secondary endpoint	OS (Overall survival)	the time from the date of randomization to the date of death due to any cause
	Other secondary endpoint	NRM (Non-relapse mortality)	the time from date of randomization to date of death not preceded by underlying disease relapse/recurrence
	Other secondary endpoint	Reduction of daily steroid dose	proportion of patients with ≥50% reduction in daily corticosteroid dose at Cycle 7 Day1
	Other secondary endpoint	Tapering of all steroids	proportion of patients successfully tapered off all corticosteroids at Cycle 7 Day 1
	Other secondary endpoint	MR (Malignancy Relapse/Recurrence)	the time from date of randomization to hematologic malignancy relapse/recurrence (only patients with underlying hematologic malignant disease)
Database lock	Interim analysis Oct 2019 and the primary analysis 08-May-2020		
Results and Analysis			
Analysis description	**Interim Analysis** (based on 196 patients i.e., 60.5% of the targeted 324 patients)		
Analysis population	FAS (Interim Analysis) = ITT		
Effect estimate per comparison	Treatment group	**Ruxolotinib**	**BAT**
	Number of subjects	97	99
			Odds ratio (RUX/BAT)
	ORR at C7 D1 (95% CI)	50.5	26.3
	Stratified CMH test	(40.2, 60.8)	(17.9, 36.1)
			HR

	FFS (median/ months) (95% CI)	NE	5.7	0.370 (0.268, 0.510)
	Log Rank test			p<0.001
Analysis description	**Primary Analysis** (based on all randomized 329 patients.			
Analysis population	FAS (primary analysis) = ITT			
	Treatment group	**Ruxolotinib**	**BAT**	
	Number of subjects	165	164	**Odds ratio** (RUX/BAT)
	ORR at C7 D1 (n(%);95% CI)	82 (49.7) 41.8, 57.6	42 (25.6) 19.1, 33.0	2.99 (1.86, 4.80) - p<0.0001
	TSS (by mLSS at C7 D1	24.2	11.0	2.62 (1.42, 4.82)
	Stratified CMH test	(17.9, 31.5)	(6.6, 16.8)	p=0.0011
	BOR up to C7 D1	76.4	60.4	2.17 (1.34, 3.52)
	Stratified CMH test	(69.1, 82.6)	(52.4, 67,9)	p=0.0011
	ORR C4 D1	54.5	31.1	2.77 (1.75, 4.39)
	Stratified CMH test	(46.6, 62.3)	(24.2, 38.8)	p<0.0001
Descriptive statistics and estimate variability	OS*	NE	NE	HR: 1.09 (0.65, 1.82)
	DOR (median/ months) (95% CI)	NE (20.2, NE)	6.24 (4.7, 13.3)	Include patients with CR/PR only
	NRM	27/165	22/164	
	MR	9/156	8/160	Patients with underlying malignancy only
	Reduction of Steroid dose ≥50%(%)	71.2	69.6	
	Tapering of all steroids (%)	31.4	27.8	
Notes	* Updated OS results with DCO 25 Jun 2021; HR: 0,596 (0.604, 1.512)			

Assessment of paediatric data on clinical efficacy

aGvHD In Study C2301, 5 adolescents were treated with RUX and 4 with BAT, with a median age of 15 (range 12-16) in the RUX arm. ORR was 80.0% (4/5) with RUX vs. 75.0% (3/4) with BAT. A total of 60% (3/5) patients in RUX arm and 75% (3/4) patients in BAT arm achieved a CR, and 20% (1/5) patients in RUX arm achieved a PR.

cGvHD In Study D2301, 4 adolescents were treated with RUX (youngest 13 years) and 8 with BAT. It is reassuring that responses with RUX were observed in 3 (75%, PR) out of 4 adolescents in the RUX arm. In comparison: 2 out of 8 adolescents (25%) in the BAT arm had a partial response up to Cycle 7 Day 1.

The inclusion of adolescents in REACH-2 and REACH-3 was based on an extrapolation concept as described in the EMA reflection paper (2017). The inclusion of adolescents was based on the following rationale:

- The disease manifestation and progression of GvHD as well as treatment options are similar in adolescents and adults;
- Evidence for children treated with ruxolitinib already existed;
- Given the maturation of the cytochrome P450 system, the metabolism of ruxolitinib in adolescents is similar to that of adults;
- Non-clinical toxicity studies did not identify toxicities that would be expected in adolescents.

The use of the same dose (10 mg BID) in adolescents as in adult patients was supported by safety and PK data of ruxolitinib from a Phase I study in pediatric patients with hematological malignancies (Loh et al 2015). As the exposure-response relationship was expected to be the same for adolescents and adults with GvHD, both allometric scaling and PBPK modelling approaches suggested that the dose for adolescents required to obtain similar exposure in terms of AUC and Cmax was similar to that of adults (20 mg BID in adults would constitute 16-18 mg BID in adolescents). In fact, the exposure in adolescents enrolled in REACH-2 and REACH-3 was in the range of the adult exposure. It was also shown that although BSA is a significant covariate, no clinically meaningful differences are to be expected in patients with a lower BSA such as adolescents.

Efficacy and safety were also expected to be similar between adolescents and adults with GvHD. Detrimental effects by ruxolitinib on bone formation and structure were not observed in toxicology safety studies using animals with a human equivalent age ≥ 12 years. Bone density measurement by dual energy x-ray absorptiometry (DEXA) were therefore not performed in REACH-2. In REACH-3, there was no clinically relevant change in bone mineral density in adolescents enrolled in the ruxolitinib arm. Overall, the safety profile was confirmed to be similar between adults and adolescents enrolled in REACH-2 and REACH-3 studies. No concerns with bone abnormalities were seen.

The rationale for extrapolation of adult data to adolescents has been adequately discussed based on disease manifestation, mode of action, non-clinical data, PK data and (expected) clinical efficacy and safety; and is acceptable.

In the product information, the recommended dose for paediatric patients with GvHD aged 12 years and older is the same as in adults, based on pivotal trials REACH 2 and REACH 3. The safety and efficacy of RUX have not been established in patients less than 12 years of age. The BSA dependent exposure is commented on in the PK section.

Clinical studies in special populations

Elderly

No specific dose adjustments are recommended in the product information for elderly patients.

Renal impairment

In line with the existing recommendation for other indications, no specific dose adjustments are included in the product information for patients with mild or moderate renal impairment. In patients with severe renal impairment (creatinine clearance less than 30 ml/min) the recommended starting dose GvHD patients is 5 mg twice daily (in line with existing recommendations for patients with PV). Please refer to PK section of this report for raised uncertainty regarding this recommendation. There are limited data to determine the best dosing options for patients with end-stage renal disease (ESRD) on haemodialysis. For GvHD patients the recommended starting dose with ESRD on haemodialysis is a single dose of 10 mg or two doses of 5 mg given 12 hours apart, to be administered post-dialysis and only on the day of haemodialysis. This recommendation is similar to PV patients. No data is available for dosing patients who are undergoing peritoneal dialysis or continuous venovenous haemofiltration.

Hepatic impairment

Dose adjustments have been recommended in the product information for PV and MF patients. However, in patients with GvHD, mild, moderate or severe hepatic impairment was not found to have a significant impact on any parameter in the population pharmacokinetic model. Therefore, no starting dose adjustment is recommended for GvHD patients with mild, moderate or severe hepatic impairment, including liver GvHD. Please refer to PK section of this report for raised uncertainty regarding this conclusion. It is acknowledged that for patients with GvHD liver involvement and an increase of total bilirubin to >3 x ULN, a recommendation has been included in the product information to monitor blood counts more frequently for toxicity and to consider a dose reduction by one dose level.

Pregnancy and lactation

There are no data from the use of RUX in pregnant or lactating woman. Animal studies have shown that RUX is embryotoxic and foetotoxic. Available pharmacodynamic/toxicological data in animals have shown excretion of RUX and its metabolites in milk. As a precautionary measure, the use of Rux during pregnancy and lactation is contraindicated.

2.5.5. Discussion on clinical efficacy

Design and conduct of clinical studies

The applicant has applied for a full marketing approval for the selective inhibitor of Janus Kinases (JAKs), JAK1 and JAK2, Jakavi (ruxolitinib) for the indication

> *Jakavi is indicated for the treatment of patients aged 12 years and older with acute graft versus host disease or chronic graft versus host disease who have inadequate response to corticosteroids or other systemic therapies (see section 5.1).*

Since the initially proposed indication (*Jakavi is indicated for the treatment of patients with graft versus host disease aged 12 years and older who have inadequate response to corticosteroids or other systemic therapies)* differentiation of the two clinical settings (acute and chronic GvHD) has been implemented and refence to section 5.1 of the SmPC has been added to allow complementation of the concise indication with some further relevant information for the Health care professionals using ruxolitinib.

Overall, the *eligibility criteria and exclusion criteria* were acceptable for both studies

Ruxolitinib dose - For both main studies dose and regimen were based on data from published data on preliminary efficacy and safety generated in patients with SR GvHD and from the supportive phase II study INCB 18424-271 (REACH-1) SR aGvHD. In addition, publications have shown that adolescents have similar toxicity profiles, maximum tolerated doses, and PK parameters compared to adults. A 10 mg b.i.d. dose in adolescents is expected to provide a similar exposure as 10 mg b.i.d. dosing in adults. The dose-selection seems overall to be adequate, however, exposure is dependent on BSA, please refer to PK section.

GvHD type and grade – Patients with higher grade acute and chronic GvHD were included in the pivotal trials with grading determined according to well established criteria. Patients with Grade I aGvHD or mild cGvHD were not eligible. In clinical practice these patients would rarely be offered (second or higher line) systemic treatment. Moreover, responses observed across GvHD stages did not indicate worse outcomes with lower grades. It is therefore considered acceptable not to specify the grade in the

indication. For details with regard to GvHD grade and type (acute or chronic) of the studied population, a reference to SmPC section 5.1 is included in the indication.

Line of treatment – Patients that had received more than one systemic treatment for SR GvHD were excluded, so eligible patients were in need of first- or second-line treatment for SR GvHD. The majority of aGvHD patients in Study C2301 (>93%) and cGvHD patients in Study D2301 (>54%) used corticosteroids and other prior systemic therapy for GvHD. The used definitions of steroid refractoriness are acceptable although this also referred to steroid-dependent patients, i.e., patients that were unable to taper corticosteroids but still responded to corticosteroids. It is reassuring that steroid-dependent patients were included with similar frequencies in both treatment arms (~28%). For acute GvHD, the definition of steroid dependency was stricter than currently recommended by the EBMT−NIH−CIBMTR Task Force position statement (Schoemans et al 2018). It could therefore be anticipated that a broader population of steroid-dependent aGvHD patients will be offered RUX treatment in clinical practice for which efficacy of ruxolitinib has not been established. Considering the persistent sensitivity to steroids in the steroid-dependent subgroup, it is anticipated that response rates in these patients will at least be similar compared to responses in steroid refractory patients or overall population. A cross reference in the indication to section 5.1 of the product information is provided, allowing prior and concomitant treatments in the studied population and used definitions of steroid refractoriness and dependency, to be captured in a concise indication.

Missing information - Due to exclusion criteria for both pivotal trials, efficacy in patients with overlap syndrome or transition from active aGvHD to cGvHD without tapering off corticosteroids ± CNI and any systemic treatment is unknown. Efficacy in acute or chronic GvHD after pre-emptive treatment of malignancy recurrence with donor lymphocyte infusion (DLI) and in patients who did not tolerate steroid treatment is unknown as well. This is specified in section 5.1 of the SmPC.

Protocol deviations - In REACH-2, protocol deviations were observed in approximately 86% in both treatment arms. The most common deviation category was 'other deviations' (~63% in both arms). It has been clarified that the overall pattern of the most frequent subcategories of 'other deviations' was comparable between the two treatment arms in both studies, with the exception of PD 'organ staging assessment done per investigator criteria/judgement rather Harris', where there was a lower proportion of patients with recorded PD in the ruxolitinib arm (18.8%) compared to the BAT arm (34.2%) in REACH2. However, for key time points (baseline, Day 28 and Day 56) these were reported with low and similar rates. The additional supportive analyses for both studies with a subset of the patients in the FAS who were compliant with the requirements of the clinical study protocol indicated similar results to that observed with the FAS. No other imbalances in subcategories 'other deviations' have been observed.

In REACH-3, protocol deviations were observed in 213 (64.7%) patients overall (63.0% in ruxolitinib arm and 66.5% in BAT arm). The most common deviation category was "Other deviations" (32.7% and 37.2%). Which included (ruxolitinib arm vs. BAT arm): Lee Symptom Scale not collected (13.9% vs. 17.1%); HBV and/or HCV viral load testing value missing at Cycle 1 Day 1 (9.7% vs. 7.9%); other deviations caused by COVID-19 pandemic as provided in Section 10.7.

A GCP inspection, proposed by the Rapporteur and Co-Rapporteur due to the concerns raised by the high numbers of protocol deviations, has been performed. The scope of the inspection were documentation and results concerning the consistency of reported data with source documents, protocol deviations relating to e.g., staging/grading of graft versus host disease (GvHD), response assessment and use of prohibited medication, and reporting of the data. Moreover, the compliance with ICH-GCP and applicable regulations was to be verified, in particular where it had an impact on the validity of the data.

In total, the inspection revealed twelve major findings and one critical finding. None of the findings could be directly related to the specific triggers. The only critical finding concerned the violation of patients' rights; race and ethnicity were captured in the electronic case report form (eCRF) without a sound rationale in the study protocols. This finding is considered as non-compliance to ICH-GCP and for the EU trial participants not in accordance with the GDPR.

During inspection, it was observed that a significant number of important protocol deviations (PDs) occurred due to lack of clarity of the study protocol and its amendments. The process for PD management was not robust enough and should have been improved earlier in order to enhance a timely resolution.

Mostly due to occurrence of PDs, for the REACH 2 clinical trial, 56 (36.4%) trial participants in the ruxolitinib arm and 68 (43.9%) in the BAT arm were excluded from the per protocol set. For the REACH 3 clinical trial 69 (41.8%) trial participants in the ruxolitinib arm and 72 (43.9%) trial participants in the BAT arm were excluded from the per protocol set.

Sensitivity analysis was conducted to assess the impact of some of these PDs on primary and key secondary efficacy endpoints as defined in the statistical analysis plans and reported in the respective clinical study reports. The results of these sensitivity analyses confirmed the results of the primary analysis.

mLSS - The second key secondary endpoint Total symptom score, TSS (rate of responders from baseline of the modified Lee Symptom Scale [mLSS]) was not validated prior to start of the trial and data are derived from an open-label study. The risk of bias due to this, cannot be measured or determined. Furthermore, it did not meet statistical significance during the interim analysis. The Applicant has agreed not to include the results in the SmPC, as requested by the Rapporteurs

BAT - The sponsor did not require documentation as detailed for BAT as for IMP (investigational medical products). The classification of IMP or non-IMP in EU member states is based on the local implementation in national law of Directive 2001/20/EC, which provides the definition. However, also for these, certain minimum documentation requirements apply. As the documentation requirement of the Sponsor failed to assure for non-IMP appropriate traceability and compliance, even these lower standards were not fulfilled.

In the worst scenario for the REACH-2, 12% of the trial participants and for the REACH-3, 41% of the trial participants receiving BAT, treatment compliance was not/insufficiently monitored.

The approximately 12% of patients without drug accountability for BAT in acute GvHD Study REACH-2 still allows sufficient interpretation of clinical study results, as impact is considered low.

For BAT in chronic GvHD study REACH-3, max 41% of patients had no data on drug accountability. Novartis conducted additional exploratory analyses to assess the primary efficacy response of the different types of BATs. This is possible as at any time point, patients in the BAT arm were administered one single BAT on top of the backbone therapy of corticosteroids and CNI.

Importantly, it is considered unlikely that all 41% of BAT treated patients were indeed non-compliant considering the severity of symptoms and known positive impact of the BAT treatment for patients. However, it cannot be excluded that (a proportion of the) patients in the BAT arm, or their treating physicians preferred a switch from BAT to Jakavi in this open label study as Jakavi was already prescribed off label in clinical practice and anticipated to be more active. This physician's/patient's preference, could have, in a worst-case scenario, influenced BAT treatment compliance in order to urge a cross-over to the ruxolitinib arm and thereby potentially have influenced treatment outcomes for ORR and durability of response.
ORR was measured at Cycle 7 Day 1, prior to the possibility for BAT treated patients to cross-over to

ruxolitinib. Now ORR results have been provided per BAT, and the observed ORR for each individual BAT is lower than the response rate reported in patients treated with ruxolitinib, regardless of the administration route and/or clinical setting i.e., in hospital versus at home. Importantly, the in-hospital BATs, which have high drug accountability (due to the setting), and the at-home BATs showed similar ORR per BAT. These exploratory results together are suggestive of a limited (if any) impact of the potential lower BAT compliance (worst case) in the at-home setting.
Regarding the time to event endpoints, demonstration of at least non-detriment in OS is considered necessary for registration in cGvHD, as discussed in the previous rounds of assessment.
As OS is multifactorial, for which BAT compliance is only one parameter of interest, it remains difficult to establish whether potential non-compliance in the BAT arm has substantially influenced OS outcomes. It is reassuring that with the latest OS update no sign of detriment was observed.
Finally, even if outcomes would have been comparable to BAT, ruxolitinib could still be acceptable as new treatment modality for chronic GvHD. Therefore, despite some remaining uncertainties on the impact of potential non-compliance on the efficacy results due to the exploratory nature of the additional analyses, the B/R of ruxolitinib is considered positive in chronic GvHD as well.

It is, however, acknowledged that patients with GvHD enrolled in the studies per study protocols were treated according to institutional guidelines, which represents local standard of care in real world practice. Transplanted patients are closely monitored by designated and experienced physicians working in specialized transplant units.

In summary, it can be concluded that in general the conduct of the REACH 2 and REACH 3 clinical trials was partially ICH-GCP compliant. Nevertheless, despite the deviation described with respect to ethics and patients´ rights in the inspections report, the REACH 2 and REACH 3 clinical trials were still considered to be conducted within internationally accepted ethical standards. The inspectors considered the data of the REACH 2 and REACH 3 clinical trials, as reported in the interim CSRs, to be acceptable and the impact of the use of the non-validated mLSS and the mostly lacking documentation on BAT accountability and treatment compliance on the benefit-risk balance is referred to the assessors.

According to the inspectors, the observed findings are unlikely to have a significant impact on data integrity within the inspected clinical trials. The inspection team did not identify any restrictions on the usability of the reported trial data, and therefore it was the recommendation of the inspectors that the data of the REACH-2 and REACH-3 clinical trials could be used for evaluation and assessment of the application. The Rapporteurs shares this view and therefore concludes that the findings are unlikely to have had any significant impact on the benefit-risk balance with regard to both aGvHD and cGvHD.

Endpoints - ORR at day 28 after starting treatment in aGvHD and 6-month ORR in cGvHD patients are acceptable primary endpoints. Start of additional therapy before the primary endpoint is reached is considered non-response (composite strategy) and so is missing data. By design, only the primary endpoints (ORR at day 28 for aGvHD or ORR at cycle 7 day 1 for cGvHD) cannot be affected by cross-over. Cross-over changes the interpretation of other endpoints and is handled by treatment policy. This may bias the effect compared to the situation when cross-over had not occurred. Under the assumption that RUX still has effect after BAT, the estimate will be conservative. Planned cross-over impacts the possibility to have reliable long-term efficacy data but should at least indicate no detrimental effect.

Censoring - Information with regard to censoring and competing risk for each of the endpoints was summarized on request. The definitions were largely the same for REACH2 and REACH3, except for DoR and FFS where onset of cGvHD was a competing event (in REACH2). These are more supportive than pivotal endpoints, so this is not critical for the B/R.

Statistical analysis methods - The statistical tests and analysis models were standard. Multiple testing procedures used control the type I error adequately. Of note: in aGvHD Study C2301, only the primary endpoint ORR at Day 28 and key secondary endpoint ORR at Day 56 were formally tested. In cGvHD Study D2301, the primary endpoint ORR at Cycle 7 Day 1 and key secondary endpoints FFS and Lee Symptom Score were formally tested. This D2301 study also included an interim analysis, at which the hypotheses for the first two endpoints (primary + FFS) were already rejected.

BAT subgroups – The choice of BAT, were the patient to be randomized to the BAT arm, was decided before randomisation. Results per the particular BATs have been provided. Furthermore, the BAT accountability has been described and commented on in the inspection report as reported above.

Changes during the study - Most notably in both study REACH 2 and REACH 3 a data monitoring committee was added during the trial conduct (this was late in the study conduct of REACH 2, to address the Study Steering Committee's request to be informed on the balance of safety events between treatment arms, while maintain blinding of the Study Steering Committee; and early in REACH 3). This is not considered to impact the results or interpretation of the endpoints. The analysis plans were not changed.

Efficacy data and additional analyses

Efficacy results for INC424C2301, aGvHD REACH-2

Baseline characteristics –Overall, baseline characteristics were generally comparable between treatment arms and reflect a steroid refractory/steroid dependent population in need of first- or second-line systemic treatment (after first line corticosteroids) for aGvHD.

Response rates - REACH-2 demonstrated a statistically significant higher ORR (primary endpoint) for patients in the RUX arm (62.3%) compared with those in the BAT arm (39.4%) (Odds ratio [HR] 2.64; 95% confidence interval [CI] 1.65, 4.22; $p < 0.0001$), thus meeting the primary endpoint, ORR (CR+PR) at Day 28. The primary endpoint results were confirmed by an analysis using PPS with an ORR at Day 28 of 63.9% in the ruxolitinib arm and 39.1% in the BAT arm. Furthermore, the robustness of primary endpoint results was confirmed by sensitivity analysis performed without stratification factor using the Fisher's exact method with an ORR at Day 28 (Odds ratio: 2.55; 95% CI: 1.61, 4.03).

Improvements in signs and symptoms were observed across all organs involved for aGvHD (skin, liver, lower GI and upper GI) at Day 28; these improvements were more pronounced in the ruxolitinib arm than in the BAT arm.

Cross over population - Among the 49 patients who crossed over to RUX treatment between Day 28 and Week 24, the ORR at Crossover Day 28 was consistent with the primary analysis in the RUX arm: 67.3% (95% CI: 52.5, 80.1). The CR rate was higher compared with the primary analysis: 46.9%.

The key secondary endpoint, durable ORR at Day 56 (proportion of patients that achieved response at Day 28 and maintained their response up to Day 56) showed statistically significant higher response rate in the RUX arm (39.6%) compared to the BAT arm (21.9%) (Odds ratio of 2.38, 95% CI: 1.43, 3.94; p=0.0005). Analysis of durable ORR at Day 56 using PPS was consistent with that observed using FAS.

Time-dependent endpoints - The magnitude of ORR and durability of response indicate activity of RUX. No nominally statistically significant differences in OS, EFS or NRM were observed, but results are impacted by the large proportion of cross-over from BAT to RUX (31.6%). It is acknowledged that results suggest no detrimental effect either and median OS and EFS were still numerically longer in the

RUX arm (primary analysis OS: 11.14 vs. 6.47 months with BAT; EFS: 8.28 vs. 4.17 months). Given the high ORR in RUX after BAT (see cross-over population), it also seems reasonable to assume that RUX has beneficial effects after BAT. For FFS, where addition of new systemic aGvHD treatment was defined as event, results obtained were numerically in favour of RUX (4.99 months vs. 1.02 months with BAT; HR: 0.46, 95% CI: 0.35, 0.60), but this endpoint was not part of the confirmatory testing strategy. The favourable effect on FFS is largely driven by less additional systemic therapy for aGvHD in the ruxolitinib arm.

Subgroups - Responses were observed across all included aGvHD severities, all categories of SR refractoriness/dependency and most combinations of prior therapies (steroid only, steroid + CNI, and most subgroups with prior steroid + CNI + other systemic treatment). In some subgroups, firm conclusions regarding efficacy cannot be drawn. considering the limited number of patients in these respective subgroups. For elderly, OR point estimate approximates to 1 suggesting similar efficacy of RUX and BAT. Even if this would be the case, RUX might still offer a valuable (registered and well characterized) treatment option that does not seem to come with additional risks compared to BAT but provides an alternative option with a different safety profile.

cGvHD - A higher proportion of patients in the RUX arm developed chronic GvHD (29.2% vs. 18.7% secondary analysis, 24.7% vs. 16.8% primary analysis) in the context of similar/better competing events (MR and NMR) for RUX. It is acknowledged that this primarily concerned lower grade events, and RUX treatment was not continued for cGvHD or overlap syndrome in the clinical trial.

Supportive study 271

Single arm, open-label Phase II study 271 provided supportive data in SR acute GvHD. The study included 71 patients with Grade II-IV SR acute GvHD that had received corticosteroids alone or in combination with 1 or more additional agents as first-line therapy for aGVHD. RUX treatment started at a lower 5 mg b.i.d. dose. Results for the primary endpoint ORR at day 28 (56.3%) were in line with results for RUX in the confirmatory Phase 3 trial (62.3%). Median OS was lower (232 days, ~7.6 months) than observed in the pivotal trial (11.14 months), but the RUX dose in this supportive trial was lower and a non-detrimental effect vs. BAT has been confirmed in the Phase 3 Study REACH-2 (C2301).

Efficacy results for the Pivotal study, INC424D2301, cGvHD REACH-3

Baseline characteristics –Overall, baseline characteristics were generally comparable between treatment arms and reflect a steroid refractory/steroid-dependent population in need of first- or second-line systemic treatment (after first line corticosteroids) for cGvHD.

Responses - At the time of the Primary Analysis, the median duration of study follow-up (from randomization to the data cut-off date) was approximately 13 months.

The proportion of patients achieving response (CR or PR) at the interim analysis at Cycle 7 Day 1 in the RUX arm (50.5%) compared to the BAT arm (26.3%) was statistically significant, with an odds ratio of 2.98 (95% CI: 1.62, 5.48) in favour of the RUX arm. The result was maintained at the primary analysis, 49.7% in the RUX arm and 25.6% in the BAT arm; odds ratio of 2.99 (95%CI: 1.86, 4.80) ($p<0.0001$). The primary endpoint results at the PA using FAS, was confirmed for the supportive analysis using PPS.

CR rates were low: 6.7% vs. 3% (RUX vs BAT). While a CR would provide most benefit to patients, it is acknowledged that the patient reported outcome (PRO) modified Lee symptom score (response rate 24.2% vs. 11%) could indicate clinical benefit in cGvHD as well, although interpretation of PRO data is hampered by the open-label study design.

Results favouring the RUX arm was seen across majority of pre-defined baseline characteristics and prognostic subgroups including gender, race, cGvHD severity (moderate/severe) at randomization, and regardless of prior diagnosis of aGvHD (odds ratio >1 for ORR at Cycle 7 Day 1). Since heterogeneity in allowed previous as well as concomitant treatment and prophylaxis regimes for GvHD makes it difficult to interpret these subgroup results. For REACH-3, only 14 patients (10 in RUX arm vs. 4 in BAT arm) had been enrolled with multiple prior therapies, and consequently ORR subgroup analyses showed large confidence intervals. It is acknowledged that it is not possible to draw firm conclusions due to the small numbers.

Cross over population - The ORR at Cycle 7 Day 1 for the cross-over population was slightly lower (39.3%) compared to the overall study population at Cycle 7 D1 (49.7%) although results became more similar with longer follow-up, 47.8% (at DCO 25-Jun-2021 used for the later OS analysis performed). These results suggest that responses can still be obtained for SR-cGvHD patients who failed treatment with BAT, to an (almost) similar level as in patients who received RUX for SR-cGvHD at an earlier stage.

The primary key secondary endpoint, FFS, showed a statistically significant difference, at the IA (N=196), between the treatment arms, with a HR of 0.315 (95% CI: 0.205, 0.486) (stratified Cox model using cGvHD severity at randomization as strata, p-value was <0.0001). Therese results were confirmed at the Primary Analysis (HR of 0.370 (95% CI: 0.268, 0.510), p<0.0001). The gain in FFS was driven by less change of therapy in the ruxolitinib arm and in lesser extent the other components of this composite endpoint, i.e., (i) relapse or recurrence of underlying disease or death due to underlying disease and (ii) non-relapse mortality. Of note: no competing risks are defined for FFS in this study, facilitating interpretation of the data. NMR and MR, both each other's competing risk, had the similar cumulative incidence, hence gain in the one was not at the cost of the other.

The second key secondary endpoint, improvement in mLSS score at Cycle 7 Day 1 did not meet statistical significance at the Interim Analysis. However, the rate of responders with improvement of ≥ 7 points in TSS at Cycle 7 Day 1, from baseline reached statistical significance, with 24.2% responders in the ruxolitinib arm and 11.0% in the BAT arm (odds ratio of 2.62 (95% CI: 1.42, 4.82) (one-sided p=0.0011) in favor of the RUX arm.

Time-dependent endpoints-. At the primary data cut-off 08 May 2020, OS and NRM medians were not reached, and results showed overlapping KM curves. For OS, there was a slightly higher number of deaths (n=31, 18.8%) in the RUX arm vs. BAT (n=27 deaths, 16.4%) in the first 6 months of treatment, and HR of 1.086 (95% CI: 0.648, 1.820). Updated OS data (25 JUN 2021) is provided, and the results remain immature, as is expected in the setting of cGvHD. The updated HR for OS shows a HR just below 1 (0.956 vs. 1.086 with previous analysis) with a CI still crossing 1. The updated results for primary cause of death did not show any new safety signals compared to the primary analysis results. Interpretation of OS analysis is hampered by cross over from the BAT to the RUX arm. A detrimental effect on OS of ruxolitinib cannot be concluded based on available immature data. Final OS data should be submitted post-marketing.

2.5.6. Conclusions on the clinical efficacy

Both pivotal studies, aGvHD and cGvHD, met their primary endpoints, showing statistical significance for the RUX arm compared to the BAT arm in terms of ORR and spared the use of other immunosuppressive treatments. This is considered clinical benefit per se, as it is anticipated to be associated with a lower symptom burden. Despite ORR not being an established trial-level surrogate for long term outcomes, efficacy is considered established.

2.6. Clinical safety

Introduction

All patients who received at least one dose of study treatment were included in the safety analysis (Safety Set).

Overview of clinical studies that contributed to safety data

Population	Studies	Study description	Dose schedule	Number of patients enrolled/treated	Data-cut-off date
Acute GvHD	CINC424C2301 (REACH 2)	A phase III randomized open-label multi-center study of ruxolitinib versus best available therapy in patients with corticosteroid-refractory acute GvHD after allogeneic stem cell transplantation	10 mg b.i.d.	309/302 Rux: 154/152 BAT: 155/150 (49 cross-over) Total number of patients exposed to Rux: 201 (51.8 PTY)	06-Jan-2020 (Secondary analysis)
	CINCB018424_271 (REACH 1)	A single-cohort, Phase II Study of ruxolitinib in combination with corticosteroids for the treatment of steroid-refractory acute GvHD	5 mg b.i.d. with possibility to increase to 10 mg b.i.d. after 3 days	71 (25.6 PTY)	05-Jun-2019 (LPLV)
Chronic GVHD	CINC424D2301 (REACH 3)	A phase III randomized open-label multi-center study of ruxolitinib vs. best available therapy in patients with corticosteroid-refractory chronic GvHD after allogenic stem cell transplantation	10 mg b.i.d.	329/323 Rux: 165/165 BAT: 164/158 (61 cross-over) Total number of patients exposed to Rux: 226 (195.1 PTY)	08-May-2020 (Primary analysis)

Data-cut-off date: when all patients completed 6 months of treatment or discontinued from study participation in Study C2301, and when all patients completed Cycle 7 Day 1 discontinued from study participation in Study D2301

Sources of safety data and safety presentation

For this presentation, no pooling of data from the studies was performed due to the differences in the study populations, study designs, duration of treatment exposure and frequency, and severity and seriousness of AE data observed across the studies during the entire treatment period.

The study population in acute GvHD and chronic GvHD are different. They differ with regard to aspects of the underlying pathophysiology, pathology, clinical manifestations, timing of transplant and disease management.

In addition, the duration of ruxolitinib treatment exposure was longer in chronic GvHD (median exposure of 41.4 weeks in REACH-3 study) compared to acute GvHD (median exposure of 8.9 weeks in REACH-2, 6.6 weeks in REACH-1).

Furthermore, the frequency and severity of adverse events was higher in REACH-2 study, compared to REACH-3 in both ruxolitinib and comparator arms, as expected in view of the recent transplant procedure and its complications in acute GvHD population. In addition, due to the differences in study design (single arm and randomized controlled study) and differences in the frequency and severity of AEs, REACH-1 and REACH-2 studies were not pooled.

Patient exposure

Duration of exposure to ruxolitinib in acute and in chronic GvHD (Safety set)

	Acute GvHD		Chronic GvHD
	Study C2301 Rux overall [1] N=201	Study 271 Rux N=71	Study D2301 Rux overall [1] N=226
Duration of exposure (weeks)			
Mean (SD)	13.4 (12.60)	18.8 (26.24)	45.1 (30.95)
Median	8.9	6.6	41.4
Q1-Q3	3.3-23.9	2.1-33.9	17.4-68.1
Min - Max	0.3-66.1	0.6-115.9	0.7-127.3
Duration of exposure categories -n (%)			
≤4 weeks	57 (28.4)	29 (40.8)	11 (4.9)
>4 to 8 weeks	38 (18.9)	14 (19.7)	12 (5.3)
>8 to 12 weeks	27 (13.4)	2 (2.8)	16 (7.1)
>12 to 16 weeks	10 (5.0)	4 (5.6)	12 (5.3)
>16 to 20 weeks	7 (3.5)	2 (2.8)	9 (4.0)
>20 to 24 weeks	23 (11.4)	2 (2.8)	11 (4.9)
>24 to 36 weeks	27 (13.4)	2 (2.8)	29 (12.8)
>36 to 48 weeks	8 (4.0)	8 (11.3)	29 (12.8)
> 48 weeks	4 (2.0)	8 (11.3)	97 (42.9)
Patient-Treatment-Years	51.8	25.6	195.1

[1] Includes the total number of patients at the time of data cut-off including the patients after cross-over.
Patient Treatment-Years (PTY) is the sum of each patient's treatment exposure in years

Exposure to BAT

Acute GvHD

Duration of exposure to BAT varied widely. The overall exposure of BAT (total exposure of BAT options used in the randomized treatment period) was 19.0 patient years (the median exposure to BAT was 29 days; range: 1.0 to 188.0). The BAT administered the longest was everolimus, followed by extracorporeal photopheresis (ECP).

Chronic GvHD

Also, in cGvHD the duration of exposure to the different regimen of BAT varied widely. The overall exposure to BAT was 24.1 weeks (range: 0.6 to 108.4) (total exposure of BAT options used in the randomized treatment period). The BAT administered the longest was ECP, followed by mycophenolate mofetil (MMF) The total exposure to ECP was 432.7 patient years.

Post-marketing exposure

Ruxolitinib was first approved for treatment of MF on 16-Nov-2011. The estimated cumulative post-marketing exposure until the end of reporting period of recent PSUR (with data cut-off of 22-Feb-2020) was estimated to be 152,580 PTY.

Due to the differences in study populations in acute GvHD and chronic GvHD the clinical manifestation and time elapsed since transplant, the extent of exposure is markedly different between the diagnoses. This accounts for both the RUX arm and for the BAT arm.

Dosage

RUX - In study REACH-2, the mean RDI up to Days 28 and 56 was 91.2% and 86.7%, with majority of patients receiving >90 to 110% of their assigned doses in both epochs. The total daily dose was 20 mg in majority of patients up to both Day 28 and Day 56 (99.3%) the majority of patients received a RUX

daily dose of 20 mg during the randomised treatment period (99.3%) and during the crossover period (91.8%).

In the supportive REACH-1 (study 271), per protocol the dosing was lower, 5 mg bid with the possibility to increase towards 10 mg b.i.d. The median daily dose received was 10.2 mg (5.1-19.7).

In study REACH-3, mean RDI (relative dose intensity) for RUX was 87.4% up to Cycle 7 Day 1 and was 83.4% for the Main treatment period. During the Cross-over period, mean (SD) dose intensity for the 61 patients was 17.8 (3.53) mg/day and mean RDI for RUX was 88.8%.

BAT - The safety sets with respect to the BAT arm for study REACH-2 functioning as the comparator set for ruxolitinib consist of 9 subsets, also there were 17 patients receiving more than one BAT treatment either simultaneously or replacing BAT 1. The safety sets with respect to the BAT arm for study REACH-3 consist of 10 subsets, also there were 16 patients receiving more than one BAT treatment either simultaneously or replacing the first BAT.

Dose adjustments for RUX

aGVHD: Majority of patients required at least one dose change or interruption of RUX (82.9%), and the primary reason was AEs (57.2%). Other frequent reasons for dose change or interruption were protocol requirements (39.5%), or dose tapering (28.9%), or by physician decision (19.1%). RUX dose interruptions occurred in 46.1% patients. In the Cross-over treatment period, all 49 patients had a dose change, and 63.3% patients had a dose interruption. Physician decision (91.8%) was the major reason for dose changes, followed by dose tapering (69.4%), AEs (61.2%), and per protocol requirements (46.9%). The CRFs did not collect the reason for the physician decision and as such the reason for dose change or dose interruption cannot be analysed for approximately 20% of the study population. Dose reescalation occurred in 38.8% patients, which was allowed as per protocol. Majority of patients (79.3%) received only one BAT post-randomisation; 18% of patients received two lines of BAT and 2.7% patients received more than 2 lines of BAT. The most common initial BAT administered to 27.3% patients was extracorporeal photopheresis (ECP). Other frequently administered BAT was mycophenolate mofetil (MMF) (16.7%) and etanercept (14.7%).

cGvHD: Up to the data cut-off date during the Main treatment period, the majority of patients required at least one dose reduction or interruption of RUX (74.5%), and the primary reason was due to AEs (47.3%). Other frequent reasons for dose change or interruption were protocol requirements (25.5%), physician decision (23.0%) and dose tapering (13.9%). RUX dose changes occurred in 69.1% patients and dose interruptions occurred in 43.0% patients.

During the Crossover treatment period, 52.5% patients required at least one dose reduction or interruption of RUX, 47.5% had a dose change, and 31.1% patients had a dose interruption. AEs (31.1%) were the major reason for dose changes, followed by physician decision (19.7%). Dose re-escalation occurred in 9.8% patients, which was allowed as per protocol. The majority of patients (80.4%) received only one BAT post-randomisation; 17.7% of patients received two lines of BAT and 1.9% patients received more than 2 lines of BAT. The most common initial BAT, administered to 55 (34.8%) patients was ECP, followed by MMF (35, 22.2%) and ibrutinib (27, 17.1%).

The comparator arms in REACH-2 study and REACH-3 study are presented below.

REACH-2

Table 14.3-1.3 (Page 1 of 1)
Initial BAT and number of BATs
Safety Set

	BAT N=150
Initial BAT -n (%)	
Extracorporeal photopheresis (ECP)	41 (27.3)
Mycophenolate mofetil (MMF)	25 (16.7)
Etanercept	22 (14.7)
Anti-thymocyte globulin (ATG)	20 (13.3)
Infliximab	17 (11.3)
Mesenchymal stromal cells (MSC)	15 (10.0)
Low-dose Methotrexate (MTX)	5 (3.3)
Sirolimus	3 (2.0)
Everolimus	2 (1.3)
Number of BAT -n (%)	
1	119 (79.3)
2	27 (18.0)
>2	4 (2.7)

CINC424C2301-Secondary CSR

REACH 3

Table 14.3-1.3 (Page 1 of 1)
Initial BAT treatment and number of BATs started before Cycle 7 Day 1
Safety set

	BAT N=158
Initial BAT treatment -n (%)	
Extracorporeal photopheresis (ECP)	55 (34.8)
Mycophenolate mofetil (MMF)	35 (22.2)
Ibrutinib	27 (17.1)
Low-dose Methotrexate (MTX)	10 (6.3)
Imatinib	8 (5.1)
Sirolimus	7 (4.4)
Rituximab	6 (3.8)
Everolimus	5 (3.2)
Infliximab	5 (3.2)
Number of BAT treatments started before Cycle 7 Day 1 -n (%)	
1	127 (80.4)
2	28 (17.7)
>2	3 (1.9)

CINC424D2301- Primary CSR

Patient disposition in Study REACH-2 and REACH-1 (aGvHD)

In addition, 49 patients randomized to BAT, crossed over to ruxolitinib treatment, whereof 36 (73.5%) patients discontinued crossover treatment period, the most common reason for discontinuation was AEs (in 24.5% of patients crossed over of whom 27 (55.1%) entered long term follow-up phase.

Patient disposition-Study REACH-2 (FAS)

	Rux N=154 n (%)	BAT N=155 n (%)	All Patients N=309 n (%)
Patients randomized			
Treated	152 (98.7)	150 (96.8)	302 (97.7)
Not treated	2 (1.3)	5 (3.2)	7 (2.3)
Treatment ongoing[1]	3 (1.9)	0	3 (1.0)
Completed treatment period[2]	35 (22.7)	21 (13.5)	56 (18.1)
Discontinued from treatment period	116 (75.3)	134 (86.5)	250 (80.9)
Reason for discontinuation			
Lack of efficacy	32 (20.8)	69 (44.5)	101 (32.7)
Adverse event	27 (17.5)	6 (3.9)	33 (10.7)
Death	25 (16.2)	21 (13.5)	46 (14.9)
Failure to meet protocol continuation criteria	12 (7.8)	9 (5.8)	21 (6.8)
Disease relapse	8 (5.2)	13 (8.4)	21 (6.8)
Physician decision	6 (3.9)	9 (5.8)	15 (4.9)
Patient/guardian decision	4 (2.6)	6 (3.9)	10 (3.2)
Graft loss	2 (1.3)	0	2 (0.6)
Lost to follow-up	0	0	0
Pregnancy	0	0	0
Protocol deviation	0	0	0
Study terminated by sponsor	0	0	0
Technical problems	0	1 (0.6)	1 (0.3)
Continued into next phase at the end of randomized treatment			
Crossover treatment	0	49 (31.6)	49 (15.9)
Entered long-term follow-up[3]	96 (62.3)	51 (32.9)	147 (47.6)

[1]Ongoing at the time of the data cut-off date 06-Jan-2020, [2]Treatment period: from Day 1 until the dosing schedule for ruxolitinib or BAT is complete or patients met any criteria for discontinuation, [3]Long term follow-up: from EoT to Month 24

Patient disposition - Study 271 REACH-1

At the time of data cut-off (05-Jun-2019), 3 patients (4.2%) who were still receiving RUX, were transferred to commercial product and 68 patients (95.8%) discontinued the study.

Patient disposition in Study REACH3 (cGvHD)

Patient disposition - REACH-3 (FAS)

	Rux N=165 n (%)	BAT N=164 n (%)	All Patients N=329 n (%)
Patients randomized			
Treated	165 (100)	158 (96.3)	323 (98.2)
Not treated	0	6 (3.7)	6 (1.8)
Treatment ongoing[1]	83 (50.3)	42 (25.6)	125 (38.0)
Discontinued from treatment period	82 (49.7)	122 (74.4)	204 (62.0)
Reason for discontinuation			
Adverse events	28 (17.0)	8 (4.9)	36 (10.9)
Lack of efficacy	24 (14.5)	70 (42.7)	94 (28.6)
Disease relapse	9 (5.5)	7 (4.3)	16 (4.9)
Death	8 (4.8)	7 (4.3)	15 (4.6)
Failure to meet protocol continuation criteria	4 (2.4)	5 (3.0)	9 (2.7)
Physician decision	4 (2.4)	14 (8.5)	18 (5.5)
Patient/guardian decision	4 (2.4)	11 (6.7)	15 (4.6)
Lost to follow-up	1 (0.6)	0	1 (0.3)
Continued into next phase at the end of randomized treatment			
Crossover treatment	0	61 (37.2)	61 (18.5)
Entered survival follow-up[2]	61 (37.0)	37 (22.6)	98 (29.8)

[1]Ongoing at the time of the data cut-off date 08-May-2020, [2] Survival follow-up: followed approximately every 3 months by telephone call for survival and reporting of new cGvHD therapies until 39 cycles completed.

There were 61 (37.2%) patients initially randomized to BAT that crossed over to ruxolitinib treatment. At DCO, 15 (24.6%) patients had discontinued the crossover treatment period with the most common reasons AEs and lack of efficacy each in 4 (6.6%) patients.

Adverse events

REACH-2

AEs considered related to the study treatment (all grades/grade ≥3) are considerably more frequent in the RUX arm compared to the BAT arm, however, for SAEs the frequencies are more or less similar.

AEs leading to treatment discontinuations and dose adjustments/interruptions was 11.32% and 36.8% in the RUX arm and 4.0% and 9.3% in the BAT arm, displaying a large difference with respect to tolerability between the study arms in aGvHD up to Day 28 (primary endpoint).

Overview of adverse events up to Day 28 visit (Safety set)

	RUX N=152		BAT N=150	
Category	All grades n (%)	Grade ≥3 n (%)	All grades n (%)	Grade ≥3 n (%)
Adverse events	**146 (96.1)**	**119 (78.3)**	**142 (94.7)**	**119 (79.3)**
Study Treatment-related	79 (52.0)	62 (40.8)	43 (28.7)	30 (20.0)
SAEs	57 (37.5)	55 (36.2)	51 (34.0)	47 (31.3)
Study Treatment-related	17 (11.2)	17 (11.2)	12 (8.0)	11 (7.3)
Fatal SAEs	12 (7.9)	12 (7.9)	18 (12.0)	18 (12.0)
Study Treatment-related	2 (1.3)	2 (1.3)	3 (2.0)	3 (2.0)
AEs leading to discontinuation	17 (11.2)	17 (11.2)	6 (4.0)	6 (4.0)
Study Treatment-related	10 (6.6)	10 (6.6)	2 (1.3)	2 (1.3)
AEs leading to study treatment dose adjustment/interruption	56 (36.8)	51 (33.6)	14 (9.3)	8 (5.3)
AEs requiring additional therapy	139 (91.4)	108 (71.1)	131 (87.3)	104 (69.3)

Numbers (n) represent counts of patients.
Treatment-related = Yes to investigational treatment, or to both (investigational treatment and CNI/corticosteroid) and/or indistinguishable.
MedDRA version 22.1, CTCAE version 4.03.

Ruxolitinib treatment up to data cut-off (including cross-over period) show that a majority of the patients (98.5%) experienced at least one AE and 66.2% of patients had AEs (all grades) related to study treatment. SAEs were reported in 69.2% of patients and treatment related SAEs were reported in 26.9% of patients. Fatal SAEs, AEs leading to discontinuation, and AEs leading to dose adjustment/interruption occurred in 23.9% of patients, 29.4% of patients, and 53.7% of patients, respectively.

REACH-3

Overview of adverse events up to Cycle 7 Day 1 (Safety set)

Category	RUX N=165 All grades n (%)	RUX N=165 Grade ≥ 3 n (%)	BAT N=158 All grades n (%)	BAT N=158 Grade ≥ 3 n (%)
Adverse events	161 (97.6)	94 (57.0)	145 (91.8)	91 (57.6)
Treatment-related	104 (63.0)	56 (33.9)	45 (28.5)	23 (14.6)
SAEs	55 (33.3)	49 (29.7)	58 (36.7)	53 (33.5)
Treatment-related	27 (16.4)	25 (15.2)	16 (10.1)	12 (7.6)
Fatal SAEs	12 (7.3)	12 (7.3)	8 (5.1)	8 (5.1)
Treatment-related	7 (4.2)	7 (4.2)	4 (2.5)	4 (2.5)
AEs leading to discontinuation	27 (16.4)	20 (12.1)	11 (7.0)	8 (5.1)
Treatment-related	16 (9.7)	12 (7.3)	6 (3.8)	5 (3.2)
AEs leading to dose adjustment/interruption	62 (37.6)	38 (23.0)	26 (16.5)	12 (7.6)
AEs requiring additional therapy	138 (83.6)	72 (43.6)	131 (82.9)	74 (46.8)

Numbers (n) represent counts of subjects.
A subject with multiple severity grades for an AE is only counted under the maximum grade.
MedDRA version 23.0, CTCAE version 4.03.

Ruxolitinib treatment up to data cut-off (including cross-over period) display similar rates of the different AE categories as reported for the period up to cycle 7, Day1.

Overview of adverse events in acute and chronic GvHD (Safety set) (at time of DCO including cross-over)

Category	Study C2301 Rux overall[1] N=201 All grades n (%)	Study C2301 Rux overall[1] N=201 Grade ≥3 n (%)	Study C271 Rux N=71 All grades n (%)	Study C271 Rux N=71 Grade ≥3 n (%)	Study D2301 Rux overall[1] N=226 All grades n (%)	Study D2301 Rux overall[1] N=226 Grade ≥3 n (%)
Adverse events	198 (98.5)	183 (91.0)	71 (100)	69 (97.2)	214 (94.7)	139 (61.5)
Treatment-related	133 (66.2)	113 (56.2)	54 (76.1)	46 (64.8)	136 (60.2)	82 (36.3)
SAEs	139 (69.2)	134 (66.7)	59 (83.1)	55 (77.5)	94 (41.6)	86 (38.1)
Treatment-related	54 (26.9)	52 (25.9)	19 (26.8)	17 (23.9)	43 (19.0)	42 (18.6)
Fatal SAEs	48 (23.9)	48 (23.9)	28 (39.4)	28 (39.4)	17 (7.5)	17 (7.5)
Treatment-related	14 (7.0)	14 (7.0)	2 (2.8)	2 (2.8)	8 (3.5)	8 (3.5)
AEs leading to discontinuation	59 (29.4)	53 (26.4)	23 (32.4)	20 (28.2)	41 (18.1)	30 (13.3)
Treatment-related	35 (17.4)	32 (15.9)	7 (9.9)	5 (7.0)	24 (10.6)	18 (8.0)
AEs leading to dose adjustment/interruption	108 (53.7)	96 (47.8)	46 (64.8)	42 (59.2)	90 (39.8)	57 (25.2)
AEs requiring additional therapy	195 (97.0)	172 (85.6)	70 (98.6)	63 (88.7)	193 (85.4)	111 (49.1)

[1] includes the total number of patients at the time of data cut-off including the patients after cross-over

Adverse events by system organ class and by PT

aGvHD:

In REACH-2, the difference in incidence of all AEs (all grades and grade≥3) between the two treatment arms (RUX arm vs. BAT arm) was ≤10% with the exception of blood and lymphatic system disorders (all grades: 58.6% vs. 44.7% and grade 3/4 AEs: 47.4% vs. 34.7%).

The most frequent AEs by SOCs (ruxolitinib treatment up to data cut-off including cross-over period) were infections and infestations (78.6% of patients), blood and lymphatic system disorders (69.2% of patients).

The incidence (at least one event) of AE Gr ≥3 was 91% in the RUX overall treatment group of REACH-2 of and 97.2% in REACH-1.

There was an increase in the cumulative incidence of most AEs by SOC, between Day 28 vs overall population in acute GvHD patients receiving ruxolitinib.

Up to Day 28, the frequently observed AEs by PT (≥20% in either treatment arm) were primarily those of cytopenias, including thrombocytopenia (32.9% vs. 18.0%), and anaemia (30.3% vs. 28.0%), followed by cytomegalovirus infection (25.7% vs. 20.7%). The difference in incidence of AEs by PT in both treatment arms was ≤10%, except for thrombocytopenia which was observed in more patients in the RUX arm (32.9%) than in BAT arm (18.0%). Up to the data cut-off date, the frequently observed AEs by PT (≥20% in either treatment arm) were those of cytopenias (including anaemia [40.1% vs. 31.3%], thrombocytopenia [36.2% vs.20.0%], neutropenia [25.0% vs. 14.7%] and platelet count decreased [20.4% vs. 16.0%]), followed by cytomegalovirus infection (30.9% vs. 26.7%), oedema peripheral (24.3% vs. 21.3%), hypokalaemia (21.1% vs. 18.7%), and pyrexia (20.4% vs. 16.7%). The difference in incidence of all AEs by PT in both treatment arms was ≤10%, except for thrombocytopenia (36.2% vs. 20.0%) and neutropenia (25.0% vs. 14.7%) which were observed in more patients in the RUX arm compared to BAT arm. The incidence of Grade ≥3 AEs by PT was similar between the two treatment arms (90.1% vs 86.7%). The grade ≥3 AEs that were more frequent (≥10% difference) in RUX arm than BAT arm were anaemia (34.2% vs 24.0%), thrombocytopenia (32.9% vs 16%) and neutropenia (22.4% vs 12%). The frequent (≥ 5% in either treatment arm) non-haematological Grade ≥3 AEs by PT (RUX vs BAT) were cytomegalovirus infection (9.2% vs 12.0%), hypokalaemia (9.2% vs 11.3%), diarrhoea (6.6% vs 5.3%), hypertension (6.6% vs 4.7%), hypoalbuminemia (5.3% vs 6.7%), sepsis (8.6% vs 11.3%) and hypoglycaemia syndrome (3.3% vs 6.0%).

Hence, the most pronounced difference (>10%) in AEs between the RUX and the BAT arm up to Day 28, was seen for thrombocytopenia (all grades and grade ≥3) in aGvHD. There were also more CMV infections/reactivations in the RUX arm compared to the BAT arm, 22.4% vs 5.6% (all grades).

REACH-2 AEs by PT up to DCO

The most frequent hematological AEs by PT (≥10% of patients) were anemia (37.8%), thrombocytopenia (35.3%), and neutropenia (23.4%), platelet count decreased (18.4%), white blood cell count decrease (11.9%), and neutrophil count decrease (11.4%).

Non-hematological AEs reported in ≥15% were hypokalemia (22.4%), cytomegalovirus infection reactivation (22.4%), pyrexia (21.4%), edema peripheral (20.9%), and nausea (16.4%).

Frequent non-hematological AEs by PT (≥10% in either treatment arm) were hypertension (15.8% vs. 12.7%), pyrexia (15.8% vs. 9.5%) and ALT elevation (15.2% vs. 4.4%), blood creatinine increased (13.9% vs 4.4%), pneumonia (10.9% vs 8.9%), cough (10.3% vs 7.0%), diarrhea (10.3% vs 13.3%), and fatigue (10.3% vs 7.6%).

In line with what is a known ADR for ruxolitinib, a higher incidence of CMV infections/reactivation is recorded in aGvHD.

Study 271: Per PT, all patients experienced at least one AE and majority of patients (97.2%) had at least one AE grade ≥3. Exposure adjusted incidence of all grade AEs and grade ≥3 AEs were 11950.6 events/100 PTY and 1666.8 events /100 PTY, respectively. All (71, 100.0%) patients experienced at least one AE and majority of them (97.2%) had at least one grade ≥3 event. Anaemia, platelet count decreased, neutrophil count decreased were the most frequently reported haematological AEs, and hypokalaemia and peripheral oedema were the most frequent non-haematological AEs.

Adverse events up to Day 28 visit, regardless of study treatment relationship in at least 10% patients in either treatment arm (all grades column), by preferred term (Safety set)

	All grades	Grade ≥3	All grades	Grade ≥3
Preferred term	**n (%)**	**n (%)**	**n (%)**	**n (%)**
Number of subjects with at least one event	**145 (95.4)**	**118 (77.6)**	**140 (93.3)**	**117 (78.0)**
Thrombocytopenia	50 (32.9)	41 (27.0)	27 (18.0)	23 (15.3)
Anaemia	46 (30.3)	33 (21.7)	42 (28.0)	28 (18.7)
Cytomegalovirus infection[1]	39 (25.7)	11 (7.2)	31 (20.7)	12 (8.0)
Oedema peripheral	28 (18.4)	2 (1.3)	26 (17.3)	1 (0.7)
Platelet count decreased	26 (17.1)	22 (14.5)	21 (14.0)	20 (13.3)
Neutropenia	24 (15.8)	20 (13.2)	19 (12.7)	14 (9.3)
Hypokalaemia	20 (13.2)	9 (5.9)	25 (16.7)	9 (6.0)
Hypertension	16 (10.5)	9 (5.9)	14 (9.3)	6 (4.0)
Hypoalbuminaemia	16 (10.5)	6 (3.9)	15 (10.0)	10 (6.7)
Pyrexia	16 (10.5)	2 (1.3)	17 (11.3)	2 (1.3)
Hypomagnesaemia	15 (9.9)	0	20 (13.3)	1 (0.7)
Diarrhoea	14 (9.2)	7 (4.6)	15 (10.0)	5 (3.3)

[1]: a distinction between cytomegalovirus infection and reactivation was not made due to MedDRA limitations (CMV reactivation coded to CMV infection).
Numbers (n) represent counts of subjects.
Preferred terms are sorted within each category in descending frequency, as reported in the all grades for the RUX treatment arm.
A subject with multiple severity grades for an AE was only counted under the maximum grade.
Adverse events occurring outside the on-randomized-treatment period or after Day 31 were not summarized.
MedDRA version 22.0, CTCAE version 4.03.

cGvHD: In REACH-3, the difference in incidence of all AEs (all grades and grade ≥3) between the two treatment arms (RUX arm vs. BAT arm) was ≤10% with the exception of blood and lymphatic system disorders (41.8% vs. 22.2%, all grades) and investigations (50.3% vs. 32.3%, all grades).

In the RUX overall treatment group, the most frequent AEs by SOCs, were infections and infestations (68.6% of patients), blood and lymphatic system disorders (42.5% of patients).

Other frequent AEs by SOCs ≥30% of patients) were investigations (49.1%), metabolic and nutrition disorders (39.8%), general disorders and administration site conditions (39.4%), gastrointestinal disorders (38.5%), respiratory, thoracic and mediastinal disorders (38.1%), and musculoskeletal and connective tissue disorders (31.4%)).

There was a similar cumulative incidence for most in AEs by SOC in the RUX population up to Cycle 7 Day 1 and in the overall period, in cGvHD patients receiving ruxolitinib, however, an increase of the cumulative incidence of AEs by SOC was obvious in infections/infestations, eye disorders, musculoskeletal and connective tissue disorders and neoplasms benign/malignant.

Up to Cycle 7 Day 1, the most frequently observed AEs by PT (≥15% in either treatment arm) were anemia (29.1% vs. 12.7%), followed by hypertension (15.8% vs. 12.7%), pyrexia (15.8% vs. 9.5%) and ALT elevation (15.2% vs. 4.4%). Of note, the exposure to RUX was much longer than to BAT up to the data cut-off.

REACH-3 AEs by PT up to DCO

The most frequent hematological AEs by PT (≥10% of patients) were anemia (28.8%), neutropenia (13.3%), and thrombocytopenia (11.5%). With the exception of pyrexia (18.1%), all other non-hematological AEs were reported in ≤ 15%.

The frequencies of AEs were similar between Cycle 7 Day 1 and up to DCO were similar.

With regard to Hematological and non-hematological AEs, up to DCO for the main treatment period, the frequently observed AEs by PT (≥15% in either treatment arm) were anemia (32.1% vs. 13.9%), pyrexia (20.0% vs. 10.8%), ALT increase (17.6% vs. 4.4%), hypertension (17.6% vs. 13.3%), blood creatinine increased (15.8% vs. 4.4%), diarrhea (15.8% vs. 15.8%), and pneumonia (15.8% vs. 13.3%).

CMV infections/reactivations were seen in 5.5% (Gr≥3: 1.2%) vs. 8.2% (Gr≥3: 0%) in the RUX and the BAT arm respectively.

Adverse events (⩾ 5% of all grades in either arm) up to Cycle 7 Day 1 regardless of study treatment relationship, by preferred term (Safety set)

	RUX N=165		BAT N=158	
Preferred term	**All grades n (%)**	**Grade ≥ 3 n (%)**	**All grades n (%)**	**Grade ≥ 3 n (%)**
Number of subjects with at least one event	161 (97.6)	94 (57.0)	145 (91.8)	91 (57.6)
Anaemia	48 (29.1)	21 (12.7)	20 (12.7)	12 (7.6)
Hypertension	26 (15.8)	8 (4.8)	20 (12.7)	11 (7.0)
Pyrexia	26 (15.8)	3 (1.8)	15 (9.5)	2 (1.3)
Alanine aminotransferase increased	25 (15.2)	7 (4.2)	7 (4.4)	0
Blood creatinine increased	23 (13.9)	0	7 (4.4)	1 (0.6)
Neutropenia	18 (10.9)	14 (8.5)	8 (5.1)	6 (3.8)
Pneumonia	18 (10.9)	14 (8.5)	20 (12.7)	15 (9.5)
Thrombocytopenia	18 (10.9)	17 (10.3)	14 (8.9)	9 (5.7)
Cough	17 (10.3)	0	11 (7.0)	0
Diarrhoea	17 (10.3)	1 (0.6)	21 (13.3)	2 (1.3)
Fatigue	17 (10.3)	1 (0.6)	12 (7.6)	3 (1.9)
Platelet count decreased	17 (10.3)	8 (4.8)	9 (5.7)	7 (4.4)
Aspartate aminotransferase increased	16 (9.7)	3 (1.8)	4 (2.5)	1 (0.6)
Dyspnoea	16 (9.7)	3 (1.8)	10 (6.3)	2 (1.3)
Hypertriglyceridaemia	16 (9.7)	8 (4.8)	13 (8.2)	6 (3.8)

Gamma-glutamyltransferase increased	15 (9.1)	11 (6.7)	5 (3.2)	3 (1.9)
Nausea	15 (9.1)	0	16 (10.1)	2 (1.3)
Headache	14 (8.5)	2 (1.2)	12 (7.6)	1 (0.6)
Upper respiratory tract infection	14 (8.5)	0	13 (8.2)	2 (1.3)
Hyperglycaemia	13 (7.9)	8 (4.8)	5 (3.2)	3 (1.9)
Hypokalaemia	13 (7.9)	3 (1.8)	16 (10.1)	7 (4.4)
Blood cholesterol increased	12 (7.3)	4 (2.4)	7 (4.4)	3 (1.9)
Constipation	12 (7.3)	0	8 (5.1)	0
Oedema peripheral	12 (7.3)	1 (0.6)	14 (8.9)	0
Vomiting	12 (7.3)	0	10 (6.3)	2 (1.3)
Amylase increased	11 (6.7)	5 (3.0)	3 (1.9)	0
Back pain	11 (6.7)	1 (0.6)	11 (7.0)	0
Insomnia	11 (6.7)	0	6 (3.8)	0
Myalgia	11 (6.7)	0	5 (3.2)	0
Urinary tract infection	11 (6.7)	1 (0.6)	5 (3.2)	2 (1.3)
Arthralgia	10 (6.1)	0	8 (5.1)	0
Lipase increased	10 (6.1)	4 (2.4)	2 (1.3)	1 (0.6)
Nasopharyngitis	10 (6.1)	0	6 (3.8)	0
BK virus infection	9 (5.5)	1 (0.6)	2 (1.3)	0
Cytomegalovirus infection reactivation	9 (5.5)	2 (1.2)	13 (8.2)	0
Hypercholesterolaemia	9 (5.5)	2 (1.2)	2 (1.3)	1 (0.6)
Hyperkalaemia	9 (5.5)	3 (1.8)	4 (2.5)	1 (0.6)
Hypomagnesaemia	6 (3.6)	0	11 (7.0)	0
Tremor	6 (3.6)	0	8 (5.1)	0
Abdominal pain	2 (1.2)	2 (1.2)	8 (5.1)	0

Numbers (n) represent counts of subjects.
A subject with multiple severity grades for an AE is only counted under the maximum grade.
MedDRA version 23.0, CTCAE version 4.03.

Adverse events suspected to be related to the study drug

In REACH-2 aGvHD, the most frequent AEs by PT (all grades) suspected to be related to study treatment (≥5% patients) in the RUX arm were haematological in nature, cytopenia including thrombocytopenia (19.1%), anemia and platelet count decreased (11.2% each), neutropenia (7.2%), WBC count decreased (5.9%), neutrophil count decreased (5.3%). All other non-haematological AEs were reported in ≤ 5%. Similarly, grade ≥3 AEs suspected to be related to RUX were also haematological AEs with the most frequent PTs including thrombocytopenia (14.5%), platelet count decreased (10.5%), anaemia (7.9%), neutropenia (5.9%), WBC count decreased (5.3%), and neutrophil count decreased (5.3% each). In the BAT arm, with the exception of platelet count decrease, white blood cell count decrease, and cytomegalovirus infection (which occurred in 4.7% of patients each), all other AEs by PT (all grades) suspected to be related to study treatment occurred in < 4%. Grade ≥3 AEs suspected to be related to study treatment were infrequent (<2%) with the exception of platelet count decreased (4.7%), decreased white blood cell count and thrombocytopenia (4.0% each), and anaemia (2.0%). Up to data cut-off (including cross-over period): AEs (all grades) and grade ≥3 AEs suspected to be related to study drug were reported in 66.2% and in 56.2% of patients. Similar to the frequencies observed up to Day 28, the most frequent AEs by PT (all grades) suspected to be related to study treatment (≥5% patients) in the RUX arm were haematological AEs. Exposure adjusted incidence of all grade AEs suspected to be related to study drug were 506.1 events/100 PTY. The most frequent exposure-adjusted incident rates (per 100 PTY) of AEs (all grades) were thrombocytopenia (85.4) and anaemia (58.6). Exposure adjusted incidence of grade ≥ 3 AEs suspected to be related to study drug were 320.1 events/100 PTY with thrombocytopenia (72.7) and anaemia (43.4) being the most frequent AEs.

Study 271: AEs (all grades) and grade ≥3 AEs suspected to be related to study drug were reported in 76.1% and in 64.8% of patients, respectively. Exposure adjusted incidence of all grade AEs and grade ≥3 suspected to be related to study drug were 722.5 events/100 PTY and 383.7 events/100 PTY.

REACH-3 cGvHD**:** Up to Cycle 7 Day 1: In Study D2301, the most frequent (≥5% patients) AEs by PT (all grades) suspected to be related to study treatment in the RUX arm were haematological AEs. Up to Cycle 7 Day 1, the most frequent (in ≥ 5% patients) AEs by PT (all grades) reported as related to study treatment in the RUX arm were anaemia (21.7%), neutropenia (10.2%) and thrombocytopenia (8.4%), ALT (8.0%), and pneumonia (5.8%). The most frequent grade ≥ 3 AEs suspected to be related to RUX by PTs were also occurred with anaemia (9.7%), thrombocytopenia (7.1%), and neutropenia (6.6%). In Study D2301, RUX treatment up to data cut-off (including cross-over period) was similar to the frequencies observed up to Cycle 7 Day 1. During the Crossover treatment period, 24 (39.3%) patients experienced at least one AE suspected to be related to study treatment, of whom 12 (19.7%) patients experienced grade ≥ 3 AEs. The frequently reported incidence (≥ 10 %) of AEs by SOC suspected to be related to RUX treatment were: blood and lymphatic system disorders (23.0% with all grades and 8.2% with grade ≥ 3 in severity), infections and infestations (14.8% with all grades and 8.2 % with grade ≥ 3), and gastrointestinal disorders (11.5% with all grades and 1.6% (n=1) with grade ≥ 3) (CSR study D2301).

With adjusted by exposure, the incidence rate of all grade AEs was higher in the RUX arm than in the BAT arm (1055.9 events/100 PTY vs. 863.3 events/100 PTY) with anaemia being the most frequent (42.9 vs 23.3 events/100 PTY), followed by pyrexia (24.4 vs 17.3) and ALT elevation (21.4 vs 6.9). (CSR study D2301). When adjusted for exposure, the incidence of grade ≥ 3 AEs was higher in the BAT arm (149.0 events/100 PTY) than in the RUX arm (123.7 events/100 PTY) with as most frequent in the RUX arm anaemia (16.9 vs. 11.9), pneumonia (14.6 vs. 15.7), neutropenia (13.5 vs. 5.8), thrombocytopenia (12.8 vs. 8.9).

AEs by age

Differences were observed in the occurrence of AEs within the age subgroups (18 to 65 years old group and >65 years old group) in the RUX treatment arm of REACH 2. Grade ≥ 3 AEs were similar across the subgroups but patients with SAEs were higher the >65 years old group (82.8%) compared to the 18 to 65 years old group (66.9%). The comparison with BAT in the elderly age group shows that in Study C2301, patients older than 65 experience more grade 3 or higher AEs in RUX arm (95.2%) than in BAT (79.2%) and more were considered treatment related in the RUX arm. This was also confirmed with higher percentages SAE, AEs leading to discontinuation, AEs leading to dose adjustments and AEs requiring additional therapy (table 2-28), even though the sample size limits accurate definition of incidences. In REACH 3, in patients older than 65 experience numerical differences were noted but assessment is hampered by the limited number of subjects older than 65 years of age (18 vs 22).Serious adverse events and deaths

Deaths aGvHD

For aGvHD and cGvHD, on-treatment deaths were defined as deaths from date of first administration of randomised treatment to 30 days after the last administration of randomised treatment or end of randomised treatment period, whichever is later.

REACH-2 - Up to Day 28: Up to Day 28, there were a total of 36 on-treatment deaths (RUX 15 (9.9%) vs BAT 21 (14.0%)). The primary reason for death was reported as study indication (including complications of the disease that were attributed to aGvHD itself or to its treatments) (5.9% RUX vs 11.3% BAT).

Up to data cut-off (including cross-over period), in RUX treated patients 62 on-treatment deaths (30.8%) were reported, of which in 29 patients (14.4%), the primary reason for death was reported as study indication (including complications of the disease that were attributed to aGvHD itself or to its treatments).

Study 271: Of the 30 deaths, 25 (35.2%) were on-treatment deaths. Seven (9.9%) were due to study indication and 18 were due to other reasons. Of note, for this study, no primary reason for death was collected. Instead, all AEs with fatal outcome were summarized. Therefore, a patient can have more than one reason for death.

On-treatment deaths in acute GvHD (Safety set including cross over)

Category Preferred term	Study C2301 Rux overall[1] N=201 n (%)	Study 271 Rux N=71 n (%)
On-treatment deaths	62 (30.8)	25 (35.2)
Primary reason: Study indication	29 (14.4)	7 (9.9)
Acute graft versus host disease	29 (14.4)	0
GvHD Progression	0	7 (9.9)
Primary reason: Other	33 (16.4)	18 (25.4)
Sepsis	4 (2.0)	1 (1.4)
Respiratory failure	3 (1.5)	3 (4.2)
Multiple organ dysfunction syndrome	2 (1.0)	2 (2.8)
Cardiac arrest	2 (1.0)	1 (1.4)
Disease progression	2 (1.0)	1 (1.4)
Septic shock	2 (1.0)	0
Pulmonary haemorrhage	1 (0.5)	1 (1.4)
Acute myeloid leukaemia	1 (0.5)	0
Acute myeloid leukaemia recurrent	1 (0.5)	0
Adenovirus infection	1 (0.5)	0
Cardiopulmonary failure	1 (0.5)	0
Cerebrovascular accident	1 (0.5)	0
Coma	1 (0.5)	0
Condition aggravated	1 (0.5)	0
Death	1 (0.5)	0
General physical condition abnormal	1 (0.5)	0
Idiopathic pneumonia syndrome	1 (0.5)	0

Category Preferred term	Study C2301 Rux overall[1] N=201 n (%)	Study 271 Rux N=71 n (%)
Infection	1 (0.5)	0
Lung disorder	1 (0.5)	0
Pneumonitis	1 (0.5)	0
Renal impairment	1 (0.5)	0
Respiratory disorder	1 (0.5)	0
Respiratory tract infection	1 (0.5)	0
Viral infection	1 (0.5)	0
Pneumonia	0	2 (2.8)
Candida infection	0	1 (1.4)
Device related infection	0	1 (1.4)
Fungaemia	0	1 (1.4)
Graft versus host disease in liver	0	1 (1.4)
Hepatic failure	0	1 (1.4)
Lower gastrointestinal haemorrhage	0	1 (1.4)
Peritonitis	0	1 (1.4)
Pneumonia legionella	0	1 (1.4)
Pulseless electrical activity	0	1 (1.4)
Staphylococcal sepsis	0	1 (1.4)
Sudden death	0	1 (1.4)
Venoocclusive liver disease	0	1 (1.4)

[1] includes the total number of patients at the time of data cut-off including the patients after cross-over.
In Study 271, no primary reason for death was collected. Instead, all AEs with fatal outcome are summarized. Therefore, a patient can have more than one reason for death.
MedDRA version 23.0

Deaths cGvHD

REACH-3- Up to Cycle 7 Day 1: A total of 13 (7.9%) patients in the RUX arm and 9 (5.7%) patients in the BAT arm died during the study. Of the 13 deaths in the RUX arm, 12 were due to study indication (including complications of the disease that were attributed to cGvHD itself or to its treatments), and one (1.6%) was due to general physical health deterioration (reported more than 30 days after last dose of RUX but before end of randomisation period).

- In Study REACH 3 up to data cut-off (including cross-over period), there were 19 deaths (8.4%) in RUX treated patients. Of these, 16 reported the primary reason for death as study indication (4 additional patients since cycle 7 day 1) (Table 7).

A more detailed description of the primary cause of death depicted as "study indication" for cGVHD has been provided. An overview of corresponding fatal SAEs according to the primary cause of on-treatment death reported as study indication up to Cycle 7 Day 1 by treatment arm with suspected treatment relationship was presented. Review of the 13 fatal cases in the ruxolitinib arm showed that the status of underlying GvHD prior to / at onset of events was either partial response or stable disease in all cases. None showed GvHD progression. It is clarified that the fatal cases of pneumonia can also be attributed to primary cause of death "study indication", of which one case was not related to study treatment and 7 cases were related to RUX and/or concomitant treatment versus 2 cases of pneumonia in the BAT arm. The relation to RUX, in the majority of the cases could not be distinguished due to the concomitant administration of prednisolone, tacrolimus, cyclosporin.

On-treatment deaths in chronic GvHD (Safety set)

Category Preferred term	Study D2301 Rux overall[1] N=226 n (%)
On-treatment deaths	19 (8.4)
Primary reason: Study indication	16 (7.1)
Chronic graft versus host disease	16 (7.1)
Primary reason: Other	3 (1.3)
General physical health deterioration	1 (0.4)
Haemorrhage intracranial	1 (0.4)
Pseudomonal sepsis	1 (0.4)

[1] includes the total number of patients at the time of data cut-off including the patients after cross-over.
MedDRA version 23.0

SAEs with fatal outcome

In REACH-2 up to Day 28, SAEs with a fatal outcome was 7.9% in RUX arm and 11.3% in BAT arm. The majority of the fatal SAEs up to Day 28 were not suspected to be related to study treatment. Fatal SAEs in 3.9% in the RUX arm and 8.7% in the BAT were due to the study indication (including complications of the disease that were attributed to aGvHD itself or to its treatments). Up to DCO (including cross-over period), a total of 48 (23.9%) ruxolitinib patients had SAEs with fatal outcomes, out of which 20 (10.0%) were attributed to study indication and 28 deaths due to other causes. The most common SAEs (≥ 1%) by PT with a fatal outcome due to other causes were respiratory failure (4 patients, 2.0%), sepsis (3 patients, 1.5%), and septic shock (3 patients, 1.5%). Study treatment related SAEs with fatal outcome were reported in 14 patients.

In REACH-1, a total of 28 (39.4%) patients had SAEs with fatal outcomes out of which 9 (12.7%) were due to study indication and 19 (26.8%) were due to other causes (most frequently reported were infections).

In REACH-3, SAEs with a fatal outcome were reported in 7.3% in the RUX arm and 5.1% in the BAT arm. Also, in REACH-3, study indication was reported to be the primary cause of fatal SAEs in 6.7% patients in the RUX arm, and 3.2% patients in the BAT arm. The most common SAE(s) (≥ 1%) by PT with a fatal outcome was pneumonia (3.0%) in the RUX arm and pneumonia and septic shock (1.3% each) in the BAT arm. Study treatment related SAEs with fatal outcome were reported in 4.2% and 2.5% patients in ruxolitinib and BAT arms, respectively.

SAEs

REACH-2 A similar proportion of patients in the RUX arm (57; 37.5%) and the BAT arm (51; 34.0%) experienced an SAE up to day 28, regardless of relationship to study treatment. Grade ≥3 SAEs were reported in 36.2% of patients in the RUX arm and 31.3% in the BAT arm (Table 35). The incidence of Grade ≥3 SAEs was 36.2% in the RUX arm and 31.3% in the BAT arm. In the RUX arm, sepsis (5.3%) was the only Grade ≥3 SAE by PT observed in >5% patients. In the BAT arm, cytomegalovirus infection (3.3%) and septic shock and respiratory failure (2.7% each), were the most frequent Grade ≥3 SAE by PT. Up to the data cut-off date, SAEs were observed in 65.1% patients in the RUX group and 52.7% patients in the BAT group. The difference was driven to some extent by the SOC of infections and infestations (38.2% in the RUX arm and 30.0% in the BAT arm).

Overall, when adjusted for exposure, the incidence of SAEs (all grades) was higher in the BAT arm (427.4 per 100 PTY) than in the RUX arm (330.8 per 100 PTY). The exposure-adjusted incidence rate (per 100 PTY) of SAEs by PT higher (≥10 events difference) in the BAT arm relative to the RUX arm were (BAT vs RUX): Sepsis (38.4 vs 26.7), Cytomegalovirus infection (28.8 vs. 13.5, and Pneumonia (27.6 vs. 13.4). The exposure-adjusted incidence rate (per 100 PTY) of Grade ≥3 SAEs by PT observed

at a higher rate (≥10 events difference) in the BAT arm relative to the RUX arm (BAT vs RUX) were: Sepsis (38.4 vs 26.7), Cytomegalovirus infection (28.8 vs 13.5), Pneumonia (27.6 vs 13.4).

Serious adverse events in REACH-2, regardless of relationship to study treatment, in at least 2% patients in either treatment arm (all grades column), by preferred term (Safety set)

	RUX N=152		BAT N=150	
	All grades	Grade ≥3	All grades	Grade ≥3
Preferred term	n (%)	n (%)	n (%)	n (%)
Number of subjects with at least one event	**99 (65.1)**	**94 (61.8)**	**79 (52.7)**	**74 (49.3)**
Sepsis	12 (7.9)	12 (7.9)	11 (7.3)	11 (7.3)
Septic shock	10 (6.6)	10 (6.6)	8 (5.3)	8 (5.3)
Pyrexia	9 (5.9)	4 (2.6)	6 (4.0)	3 (2.0)
Diarrhoea	7 (4.6)	7 (4.6)	3 (2.0)	2 (1.3)
Cytomegalovirus infection	6 (3.9)	6 (3.9)	8 (5.3)	8 (5.3)
Pneumonia	6 (3.9)	6 (3.9)	8 (5.3)	8 (5.3)
Respiratory failure	6 (3.9)	6 (3.9)	6 (4.0)	6 (4.0)
Neutropenia	4 (2.6)	4 (2.6)	3 (2.0)	3 (2.0)
Acute kidney injury	3 (2.0)	3 (2.0)	4 (2.7)	4 (2.7)
Blood bilirubin increased	3 (2.0)	3 (2.0)	0	0
Cytomegalovirus colitis	3 (2.0)	3 (2.0)	0	0
Febrile neutropenia	3 (2.0)	3 (2.0)	2 (1.3)	2 (1.3)
Pancytopenia	3 (2.0)	3 (2.0)	0	0
Multiple organ dysfunction syndrome	2 (1.3)	2 (1.3)	3 (2.0)	3 (2.0)
Renal failure	2 (1.3)	1 (0.7)	3 (2.0)	2 (1.3)
Pseudomonal sepsis	3 (2.0)	3 (2.0)	0	0
Bacteraemia	1 (0.7)	1 (0.7)	4 (2.7)	3 (2.0)
Confusional state	1 (0.7)	1 (0.7)	3 (2.0)	2 (1.3)
Graft versus host disease	1 (0.7)	1 (0.7)	3 (2.0)	3 (2.0)
Platelet count decreased	0	0	3 (2.0)	3 (2.0)

Numbers (n) represent counts of subjects.
Preferred terms are sorted within each category in descending frequency, as reported in the all grades for the RUX treatment arm.
A subject with multiple severity grades for an AE was only counted under the maximum grade.
Adverse events occurring outside the on-randomized-treatment period are not summarized.
MedDRA version 22.0, CTCAE version 4.03.

SAE related to study treatment:

REACH-2: Up to Day 28, the incidence of SAEs suspected to be related to study treatment was similar between the RUX arm and BAT arm (10.5% vs. 7.3%), with as most common SOC in either treatment arm: infections and infestations (5.9% RUX vs 4.0% BAT). All other SOCs were observed in <2% patients in either treatment arm. The treatment related SOCs of cardiac disorders, general disorders and administrative site conditions, immune system disorders, hepatobiliary disorders and renal and urinary disorders were observed only in the BAT arm. Except for PT sepsis observed in 3 patients randomized to RUX arm, no other SAE by PT occurred in more than 1 or 2 patients each.

Up to the data cut-off date, the incidence of SAEs suspected to be related to study treatment was higher in the RUX arm compared to BAT arm (25.7% vs. 11.3%). Except for PT sepsis, observed in 5 patients randomized to RUX arm, no other SAE by PT occurred in more than 1 or 2 patients each. In all these 5 patients reported with sepsis, the events were attributed to both RUX and other study treatments (CNIs and/or steroids). The single SAE of graft loss in the RUX arm was observed at data cut-off. After crossover, the incidence (24.5%) and overall profile of SAEs suspected to be related to study treatment was consistent with that observed in patients randomized to RUX.

In the crossover period, 77.6% patients had SAEs and Grade ≥3 SAEs were reported in (75.5%) patients and the most common Grade ≥3 SAEs by PT were: sepsis (14.3%) and respiratory failure

(12.2%). Up to the second analysis DCO, the incidence of SAEs suspected to be related to study treatment was higher in the RUX arm compared to BAT arm (27.0% vs. 12.0%), driven by higher incidence (≥5% difference) of the SOCs in the blood and lymphatic system disorders (5.3% vs 0.7%) and infections and infestations (14.5% vs. 6.7%) in the RUX arm relative to the BAT arm. After crossover, the incidence and overall profile of SAEs suspected to be related to study treatment was consistent with that observed in ruxolitinib arm at second analysis cutoff.

Study 271: SAEs were reported in 83.1% of patients. Exposure adjusted incidence of SAEs was 487.2 events/100 PTY. The most frequent SAEs by exposure adjusted incidence were pneumonia (33.9%), sepsis (33.5%), and pyrexia (31.8%).

REACH-3

Up to Cycle 7 Day 1, a similar proportion of patients in the RUX arm (33.3%) and the BAT arm (36.7%) experienced an SAE. By PT, the frequent (≥ 2% patients in either treatment arm) SAEs (RUX vs. BAT) were: pneumonia (7.9% vs. 8.2%), pyrexia (4.8% vs. 1.9%), lower respiratory tract infection (2.4% vs. 0), and bronchopulmonary aspergillosis (1.2% vs. 2.5%). The incidence of grade ≥ 3 SAEs was 29.7% in the RUX arm and 33.5% in the BAT arm. Up to the data cut-off date for the Main treatment period, the overall SAE profile was similar to those up to Cycle 7 Day 1 with a slight increase in the incidence as expected due to longer exposure. SAEs were observed in 43.6% of patients in the RUX arm and 39.9% of patients in the BAT arm.

Overall, when adjusted for exposure, the incidence of SAEs (all grades) was higher in the BAT arm (76.6 per 100 PTY) than in the RUX arm (60.3 per 100 PTY). Similarly, the incidence of grade ≥ 3 SAEs was higher in the BAT arm (65.4 events/100 PTY) than in the RUX arm (51.6 events/100 PTY). The Exposure-adjusted incidence rate of most commonly observed (with >5 per 100 PTY) all grades SAEs (by PT) in the RUX arm and the BAT arm (RUX and BAT, respectively) were: pneumonia (13.8 and 13.7 per 100 PTY) and pyrexia (7.3 and 3.9 per 100 PTY).

SAE related to study treatment:

Up to Cycle 7 Day 1, the incidence of SAEs reported as related to study treatment was 16.4% in the RUX arm and 10.1% in the BAT arm. The most common SAE (by PT) reported as related to study treatment was pneumonia in 6.1% of patients in RUX arm and 2.5% of patients in the BAT arm, followed by pyrexia 2.4% (4 patients) in RUX arm vs. 0.6% (1 patient) in the BAT arm. Other SAEs by PT occurred in < 2 patients in each arm. Up to the data cut-off date for the Main treatment period, the incidence of SAEs reported as related to study treatment was higher in the RUX arm compared to the BAT arm (21.2% vs. 10.1%). The incidence of SAEs (by PT) with pneumonia was higher in the RUX arm (n=12, 7.3%) compared to that in the BAT arm (n=4, 2.5%). Pyrexia was observed in 5 patients (3.0%) in the RUX arm vs. 1 patient (0.6%) in the BAT arm and lower respiratory tract infection was observed in 4 patients (2.4%). Other SAE by PT occurred in no more than 2 patients in each arm.

After cross over to RUX treatment, SAEs related to study treatment were reported in 8 (13.1%) patients with all grades and in 8 (13.1%) patients with grade ≥ 3. The most common (≥ 5%) SAEs suspected to be related to study treatment by SOC (all grade, grade ≥ 3) were: infections and infestations (8.2%, 8.2%) and blood and lymphatic system disorders (6.6%, 4.9%).

AEs leading to treatment discontinuations

REACH-2 aGvHD

Up to Day 28, AE(s) leading to permanent discontinuation of study treatment were reported in 11.2% and 4.0% patients in the RUX arm and BAT arm, respectively, of whom most were of Grade ≥3 in

severity. The most common AEs (≥ 2%) by PT leading to discontinuation in RUX arm were anemia and thrombocytopenia. In the BAT arm, no single AE (by PT) occurred in more than one patient leading to discontinuation.

Up to the second DCO, AEs leading to discontinuation were observed in 27% in the RUX arm and 9.3% in the BAT arm. The majority of AEs leading to discontinuation in both arms were Grade≥3 in severity (23.0% and 8.7%). The most common (≥2%) AEs leading to discontinuation in the ruxolitinib arm were neutropenia, sepsis (2.6% each) and anemia and thrombocytopenia (2.0% each).

Study 271 AEs (all grades) leading to discontinuation were reported in 32.4% of patients and grade ≥3 AEs in 28.2% matching the frequencies in REACH-2.

REACH-3 cGvHD

Up to Cycle 7 Day 1, AEs leading to discontinuation were observed in 16.4% in the RUX arm and 7.0% in the BAT arm. The majority of the AEs were grade ≥ 3 and pneumonia was the only AE (all grades) leading to discontinuation in at least 2% patients in ruxolitinib arm (4.8%, grade ≥ 3: 3,6%). Discontinuations due to pneumonia were 1.3% in BAT arm.

Adverse events, leading to study drug discontinuation, by PT (Safety Set) (≥ 2 patients)

	Study D2301 Rux overall[1] N=226	
Preferred term	**All grades n (%)**	**Grade ≥3 n (%)**
Number of patients with at least one event	41 (18.1)	30 (13.3)
Pneumonia	9 (4.0)	7 (3.1)
Thrombocytopenia	2 (0.9)	2 (0.9)
Haemorrhage intracranial	2 (0.9)	2 (0.9)
Septic shock	2 (0.9)	2 (0.9)
Alanine aminotransferase increased	2 (0.9)	1 (0.4)
Pneumothorax	2 (0.9)	1 (0.4)
Anaemia	2 (0.9)	0

[1] includes the total number of patients at the time of data cut-off including the patients after cross-over.
Numbers (n) represent counts of patients.
A patient with multiple severity grades for an AE is only counted under the maximum grade.
MedDRA version 23.0, CTCAE version 4.03.

Up to DCO, AEs (all grades) leading to discontinuation were reported in 18.1% and grade ≥3 AEs in 13.3% of ruxolitinib treated patients.

In aGvHD most treatment discontinuations were due to hematological AEs while most treatment discontinuations in cGvHD were due to infections.

It should be noted that discontinuation criteria were outlined in protocol only for ruxolitinib treatment while discontinuation criteria of BAT treatments were left at INV's judgement per standard of care.

AEs leading to dose adjustment and/or interruption

REACH-2 aGvHD

Up to Day 28, 36.8% and 9.3% the RUX arm and BAT arm, respectively, experienced an AE leading to dose adjustment or interruption. Majority of the AEs leading to dose adjustment or interruption in both treatment arms were Grade ≥3 in severity (33.6% and 5.3%).

The higher incidence of AEs leading to dose adjustment or interruption in the RUX arm compared to the BAT arm was driven by cytopenias (Thrombocytopenia: 14.5% vs. 1.3%, Neutropenia 11.8% vs. 0.7%, Platelet count decreased 11.2% vs. 2.7%, Leukopenia (5.3% vs. 0.7%).

Up to DCO, 54.6% of patients on ruxolitinib treatment experienced AEs (all grades) leading to dose adjustment or interruption, driven by cytopenias. In BAT arm 13.3% were dose adjusted or interrupted Most of the AEs were Grade≥3 in severity in both treatment arms (50.0% and 9.3%).

In REACH-1, AEs (all grades) leading to dose adjustment or interruption were seen in 64.8% patients (grade≥3: 59.2%) and cytopenias were the most common reason.

REACH-3 cGvHD

Up to Cycle 7 Day 1, 37.6% (grade ≥3. 23.0%) in the RUX arm and 16.5% (grade ≥3: 7.6%) in the BAT arm experienced an AE leading to dose adjustment or interruption. In the RUX arm, anemia (7.9%) and blood creatinine increased (6.1%) were the most frequent AEs leading to dose adjustment or interruption (≥5%). In the BAT arm, except for pyrexia observed in 3 patients, no other PTs were observed in more than two patients.

Up to DCO for the main treatment period (including cross-over period) ruxolitinib treatment AEs lead to a dose adjustment or interruption in 43.0% (grade ≥3: 28.5%). In the BAT arm these events occurred in 18.4% (grade ≥3: 8.2%). In the RUX arm, AEs occurring ≥5% (all grades) and leading to dose adjustment or interruption were anemia (8.8%, grade≥ 3: 4.0%), and neutropenia (6.6%). Most frequent (≥2%) grade≥ 3 AEs were neutropenia (5.3%), anemia (4.0%), pneumonia (3.5%), thrombocytopenia (2.7%), and ALT increased (2.7%).

Criteria for dose adjustment/ interruption for ruxolitinib were outlined in the protocol, however with respect to BAT this was left at the Investigator's judgement.

Adverse events of special interest

AESIs are defined based on known safety data.

The aggregated PTs used to identify the AESI were based on the standardized MedDRA Queries (SMQs) /or group of PTs using MedDRA version 22.1 for REACH-2 and version 23.0 for REACH-3.

Note that the AESIs, in the RUX arm and in the BAT arm, are presented up to Day 28 in REACH-2, aGvHD and up to Cycle 7 Day 1 in REACH-3, cGvHD.

Further, the safety data (AESI) for the entire treatment period up to cut-off date for ruxolitinib treated patients (including patients originally randomized to BAT and who crossed over to ruxolitinib) is presented (for aGvHD and cGvHD) followed by exposure adjusted analysis.

Overview of adverse events of special interest in acute and chronic GvHD

Safety topic	Study C2301 Rux overall[1] N=201		Study 271 Rux N =71		Study D2301 Rux overall[1] N = 226	
	All grades n (%)	Grade ≥3 n (%)	All grades n (%)	Grade ≥3 n (%)	All grades n (%)	Grade ≥3 n (%)
Infections						
Infections excluding Tuberculosis	159 (79.1)	104 (51.7)	58 (81.7)	46 (64.8)	157 (69.5)	59 (26.1)
Opportunistic infections	67 (33.3)	26 (12.9)	15 (21.1)	7 (9.9)	26 (11.5)	7 (3.1)
CMV infection/disease	65 (32.3)	23 (11.4)	14 (19.7)	6 (8.5)	20 (8.8)	3 (1.3)
Sepsis and septic shock	51 (25.4)	44 (21.9)	16 (22.5)	15 (21.1)	8 (3.5)	8 (3.5)
Pneumonia	36 (17.9)	26 (12.9)	17 (23.9)	15 (21.1)	47 (20.8)	36 (15.9)
Urinary tract infections (UTI)	36 (17.9)	13 (6.5)	10 (14.1)	6 (8.5)	21 (9.3)	3 (1.3)
Herpes zoster	4 (2.0)	2 (1.0)	0	0	4 (1.8)	3 (1.3)
Hepatitis B reactivation	1 (0.5)	0	0	0	1 (0.4)	0
Other infections	128 (63.7)	60 (29.9)	42 (59.2)	30 (42.3)	131 (58.0)	26 (11.5)
Cytopenias						
Thrombocytopenia	107 (53.2)	94 (46.8)	44 (62.0)	38 (53.5)	45 (19.9)	33 (14.6)
Leukopenia	89 (44.3)	82 (40.8)	42 (59.2)	35 (49.3)	50 (22.1)	37 (16.4)
Erythropenia (Anemia)	78 (38.8)	64 (31.8)	46 (64.8)	36 (50.7)	65 (28.8)	31 (13.7)
Other Cytopenias	18 (9.0)	14 (7.0)	0	0	2 (0.9)	1 (0.4)
Bleeding events						
Bleeding (Haemorrhage)	74 (36.8)	25 (12.4)	35 (49.3)	14 (19.7)	26 (11.5)	6 (2.7)
GI bleeding	33 (16.4)	16 (8.0)	19 (26.8)	9 (12.7)	2 (0.9)	2 (0.9)
Bruising	21 (10.4)	2 (1.0)	10 (14.1)	1 (1.4)	9 (4.0)	0
Intracranial haemorrhage	2 (1.0)	2 (1.0)	1 (1.4)	0	2 (0.9)	2 (0.9)
Other haemorrhage events	35 (17.4)	5 (2.5)	22 (31.0)	4 (5.6)	15 (6.6)	2 (0.9)
Others						
Hypertension	27 (13.4)	11 (5.5)	16 (22.5)	10 (14.1)	34 (15.0)	13 (5.8)
Elevated Transaminases	30 (14.9)	13 (6.5)	22 (31.0)	6 (8.5)	39 (17.3)	14 (6.2)
Thromboembolic events	25 (12.4)	14 (7.0)	16 (22.5)	9 (12.7)	10 (4.4)	5 (2.2)
Lipid abnormalities	21 (10.4)	11 (5.5)	9 (12.7)	3 (4.2)	40 (17.7)	13 (5.8)
Dizziness	12 (6.0)	0	10 (14.1)	0	7 (3.1)	0
Fracture	7 (3.5)	3 (1.5)	5 (7.0)	2 (2.8)	8 (3.5)	1 (0.4)
Weight gain	3 (1.5)	0	3 (4.2)	1 (1.4)	8 (3.5)	0
Second Primary Malignancies	4 (2.0)	3 (1.5)	5 (7.0)	3 (4.2)	4 (1.8)	1 (0.4)
Non melanoma skin cancers	0	0	1 (1.4)	0	3 (1.3)	1 (0.4)

[1] includes the total number of patients at the time of data cut-off including the patients after cross-over
A patient with multiple severity grades for an AE is only counted under the maximum grade
Adverse events occurring outside the on-randomized-treatment period or after Day 31 are not summarized.
MedDRA version 23.0, CTCAE version 4.03, Case Retrieval Strategy version released 2020-08-19.

It is noted that out of the selected AESIs, growth retardation, PML and tuberculosis, there were no patients under these safety topics in any of the three studies.

In addition, there were only four patients in each pivotal study developing herpes zoster and one patient in each study with a hepatitis B reactivation.

Infections

Infection is major complication following an allo-transplant procedure. Furthermore, immunosuppressive agents given prophylactic for GvHD increase the risk of opportunistic viral and fungal pathogens as well as either a risk of reactivation of CMV or a de novo CMV infection (due to mis-match donor with respect to CMV status). Due to the transplant procedure all patients are also heavily myelosuppressed.

REACH-2 aGvHD

Up to Day 28, the proportion of patients with infections were comparable in the RUX arm (61.2%) vs. the BAT arm (58.7%).

Up to data cut-off, the exposure adjusted overall incidence of infections in the RUX arm was 681.9/100 PTY and 787.0/100 PTY in the BAT arm.

Up to DCO, ruxolitinib exposure adjusted overall incidence rate was 652.2/100 PTY. Treatment-related AEs and SAEs were noted in 24.9% and 39.8% respectively. At the time of data cut-off, in 57.7% AEs were recovered/resolved and in 34.3% not recovered/not resolved. The most frequent PTs (≥ 5%) were CMV infection reactivation (22.4%), sepsis (12.4%), pneumonia (10.4%), urinary tract infection (8.5%), CMV infection (7.5%), device related infection (7.0%), and septic shock (6.5%). The most common fatal AEs were sepsis (10 deaths), septic shock (7 deaths), pneumonia (2 deaths), and pseudomonal sepsis (2 deaths). The majority of infections occurred within the first 2 months of treatment.

Until the DCO, sepsis and septic shock was 24.3% in the RUX arm and 20.7% in the BAT arm (of which most were of Grade ≥3 in severity. Study treatment discontinuations and dose interruptions were recorded in 4.6% vs 2.0% and dose interruptions was 4.6% in the RUX arm. However, up to the Secondary analysis data cut off, the exposure-adjusted IR of PT sepsis (all grades) was lower in ruxolitinib arm (33.2 events/100 PTY) compared to BAT arm (66.1 events/100 PTY).

REACH-3 cGvHD

Up to Cycle 7 Day 1: Proportion of patients with infections were comparable in the ruxolitinib arm (62.4%) vs. the BAT arm (58.2%).

Up to data cut-off (excluding cross-over), the exposure adjusted overall incidence rates was 169.0/ 100 PTY in the ruxolitinib arm and 185.0 /100 PTY in the BAT arm.

Up to DCO, ruxolitinib exposure adjusted overall incidence rate was 161.2/100 PTY. Treatment-related AEs and SAEs were noted in 26.5% and 27.0% respectively. At the time of data cut-off, 58.0% AEs were recovered/resolved (Table 2-15). The most frequent PTs (≥5%) were pneumonia (14.6%), upper respiratory tract infection (13.3%), influenza (8%), nasopharyngitis (7.1%), bronchitis, urinary tract infection (6.6% each), conjunctivitis and CMV reactivation (5.3%).

Majority of the infections occurred within the first 4 months of treatment.

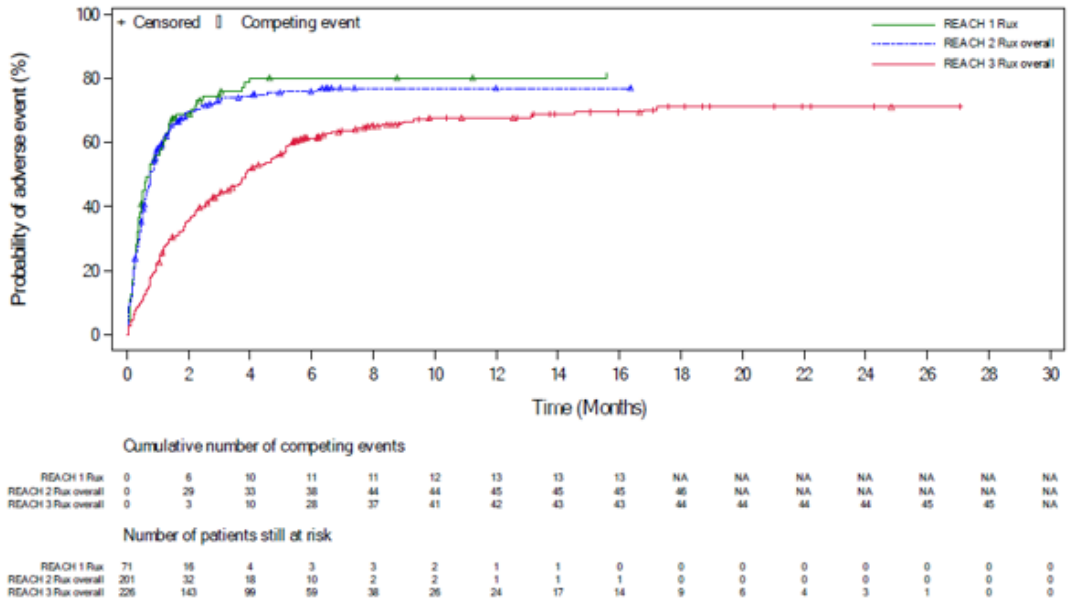
Infections excluding tuberculosis in acute and chronic GvHD

	Study C2301 Rux overall[1] N=201	Study 271 Rux N=71	Study D2301 Rux overall[1] N=226
Number of patients with at least one event, n (%)	159 (79.1)	58 (81.7)	157 (69.5)
95% CI	(72.8,84.5)	(70.7,89.9)	(63.0,75.4)
Exposure-adjusted overall incidence,n (EAIR per 100 PTY)	159 (652.2)	58 (652.6)	157 (161.2)
Maximum grade, n (%)			
Grade 3 AEs	55 (27.4)	26 (36.6)	42 (18.6)
Grade 4 AEs	23 (11.4)	20 (28.2)	5 (2.2)
Grade 5 AEs	26 (12.9)	NA[2]	12 (5.3)
Treatment-related AEs, n (%)	50 (24.9)	21 (29.6)	60 (26.5)
SAEs, n (%)	80 (39.8)	36 (50.7)	61 (27.0)
Action taken, n (%)			
Drug Withdrawn	17 (8.5)	9 (12.7)	18 (8.0)
Dose Reduced	2 (1.0)	2 (2.8)	11 (4.9)
Dose Increased	0	0	1 (0.4)
Drug Interrupted	20 (10.0)	8 (11.3)	18 (8.0)
Dose Not Changed/NA/Unknown	153 (76.1)	46 (64.8)	148 (65.5)
Missing	0	0	0
Medication or therapy taken, n (%)	150 (74.6)	56 (78.9)	139 (61.5)
AE outcome, n (%)			
Recovered/Resolved	116 (57.7)	42 (59.2)	131 (58.0)
Recovering/Resolving	11 (5.5)	5 (7.0)	16 (7.1)
Not Recovered/Not Resolved	69 (34.3)	29 (40.8)	32 (14.2)
Recovered/Resolved With Sequelae	5 (2.5)	1 (1.4)	2 (0.9)

	Study C2301 Rux overall[1] N=201	Study 271 Rux N=71	Study D2301 Rux overall[1] N=226
Fatal	26 (12.9)	10 (14.1)	12 (5.3)
Unknown	3 (1.5)	0	0
Missing	0	0	2 (0.9)

[1] All patients treated with ruxolitinib including patients who switched to ruxolitinib after cross-over.
EAIR = exposure-adjusted incidence rate: number of patients with an event divided by the corresponding sum of the exposure duration for all patients, where duration of exposure in patient-Treatment-Years (PTY) is counted up to the first qualifying event (or end of time at risk for patients without event).
[2]In REACH 1 all fatal AEs are reported as grade 4. Grade 5 is NA. Treatment-related = Yes to investigational treatment, or to both investigational treatment and non-investigational treatment (CNIs or steroids) and/or indistinguishable.
MedDRA version 23.0, CTCAE version 4.03, Case Retrieval Strategy version released 2020-08-19

Cumulative incidence of infections excluding tuberculosis in <u>acute and chronic</u> GvHD (Safety set)



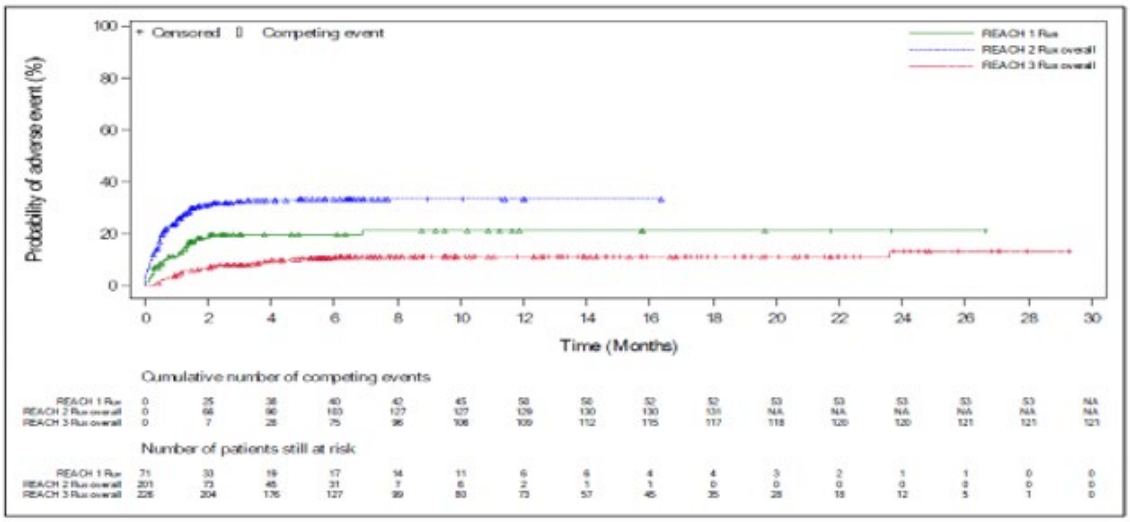
Competing risks: death or discontinuation due to any reason

Pneumonia/opportunistic infections

The proportion of patients with pneumonia during period up to Day 28 (aGvHD) or up to Cycle 7, Day 1 (cGvHD), were comparable between the treatment arms. The exposure-adjusted overall incidence (EAIR per 100 STY), was considerably higher in aGvHD compared to cGvHD.

The proportion of opportunistic infections during period up to Day 28 (aGvHD) was 27% vs 22% in the RUX vs BAT arm and for cGvHD, up to Cycle 7 Day 1, 11% vs 12% in the respective treatment arm. The exposure-adjusted overall incidence (EAIR per 100 STY), was considerably higher in aGvHD compared to cGvHD.

Cumulative incidence of opportunistic infections in acute and chronic GvHD (Safety set)



CMV infection/disease and viral other infections

REACH-2 aGvHD

Up to Day 28, CMV infections were reported in 28.3% in the RUX arm vs. 24.0% in the BAT arm and during the period up to data cut-off, in 35.5% (Grade≥3: 12.5%) in the RUX arm and 32.0% (Grade ≥3: 14.0%) in the BAT arm.

The exposure adjusted overall incidence rate (excluding the crossover period), 154.0/100 PTY in the RUX arm and 214.5 /100 PTY in the BAT arm. At the time of data cut-off, in 21.4% AEs were recovered or resolved.

REACH-3 cGvHD

Up to Cycle 7 Day 1: The proportions of patients with CMV infections were similar in the RUX arm (9.1%) vs. the BAT arm (10.8%). CMV disease (enteritis, pneumonia) was reported in 1.2% and 1.3% in the ruxolitinib arm and in the BAT arm, respectively.

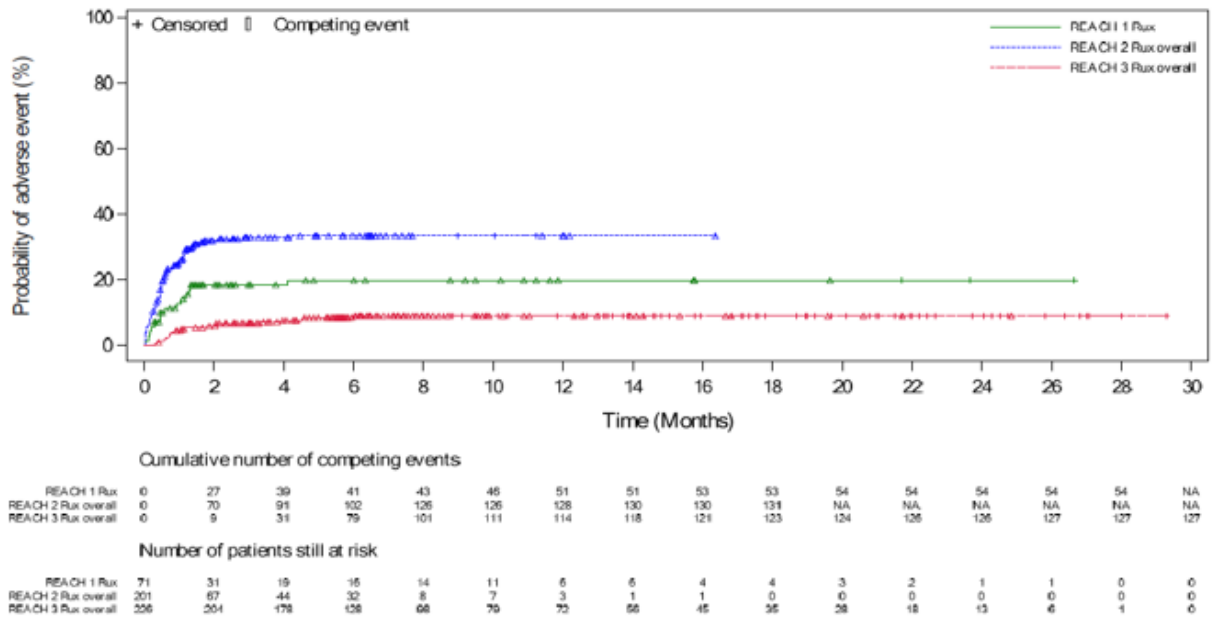
Up to data cut-off (excluding cross-over), the exposure adjusted overall incidence rates was 10.7/100 PTY in the ruxolitinib arm and 18.4 /100 PTY in the BAT arm.

CMV infection/disease in acute and chronic GvHD (Safety set)

	Study C2301 Rux overall[1] N=201	Study 271 Rux N=71	Study D2301 Rux overall[1] N=226
Number of patients with at least one event, n (%)	65 (32.3)	14 (19.7)	20 (8.8)
95% CI	(25.9,39.3)	(11.2,30.9)	(5.5,13.3)
Exposure-adjusted overall incidence,n (EAIR per 100 STY)	65 (142.3)	14 (55.5)	20 (10.0)
Maximum grade, n (%)			
Grade 3 AEs	22 (10.9)	6 (8.5)	3 (1.3)
Grade 4 AEs	1 (0.5)	0	0
Grade 5 AEs	0	NA	0
Treatment-related AEs, n (%)	14 (7.0)	4 (5.6)	7 (3.1)
SAEs, n (%)	14 (7.0)	1 (1.4)	4 (1.8)
Action taken, n (%)			
Drug Withdrawn	2 (1.0)	1 (1.4)	0
Dose Reduced	1 (0.5)	0	0
Dose Increased	0	0	0
Drug Interrupted	2 (1.0)	1 (1.4)	1 (0.4)
Dose Not Changed/NA/Unknown	63 (31.3)	12 (16.9)	20 (8.8)
Missing	0	0	0
Medication or therapy taken, n (%)	64 (31.8)	14 (19.7)	19 (8.4)
AE outcome, n (%)			
Recovered/Resolved	43 (21.4)	9 (12.7)	17 (7.5)
Recovering/Resolving	6 (3.0)	0	0
Not Recovered/Not Resolved	18 (9.0)	5 (7.0)	2 (0.9)
Recovered/Resolved With Sequelae	1 (0.5)	0	0
Fatal	0	0	0
Unknown	0	0	0
Missing	0	0	1 (0.4)

[1] All patients treated with ruxolitinib including patients who switched to ruxolitinib after cross-over.
EAIR = exposure-adjusted incidence rate: number of patients with an event divided by the corresponding sum of the exposure duration for all patients, where duration of exposure in Patient-Treatment-Years (PTY) is counted up to the first qualifying event (or end of time at risk for patients without event).
In REACH 1 all fatal AEs are reported as grade 4. Treatment-related = Yes to investigational treatment, or to both and/or indistinguishable.
MedDRA version 23.0, CTCAE version 4.03, Case Retrieval Strategy version released 2020-08-19

Cumulative incidence of CMV infections/disease in acute and chronic GvHD (Safety set)



With respect to acute vs chronic GvHD, the increased susceptibility of opportunistic infections and CMV infections/reactivation in aGvHD patients on treatment, is clearly demonstrated.

AESI CMV infections were reported for REACH-3 in 9.1% (Grade≥3: 1.8%) up to Cycle 7 Day 1 (Safety set) and approximately the same rate for the main treatment period. Furthermore, CMV infection using other PTs were also reported, in a lower frequency.

Adenovirus infections

A review of the available cases of Adenovirus infection (n=26) in patients treated with ruxolitinib for GvHD, has been provided, including cases from clinical studies REACH-2 and REACH-3, spontaneous reports and literature.

There were 11 paediatric cases and 15 cases reported in adults. All cases were either confounded with multiple immunosuppressant therapy or were poorly documented which precluded evaluation of cases. There was none case with ruxolitinib monotherapy or a case where causal association only with ruxolitinib was established or suspected. However, having considered multiple immunosuppressive treatment, occurrence or reactivation of infections, including viral, could be expected.

The available data does not allow to assess the role of ruxolitinib on the incidence of adenoviral infection in patients with SR GvHD.

Other Viral infections

Proportion of aGvHD patients with viral infections, during the period up to Day 28, were higher in the RUX arm (39.5%, Grade≥3 14.5%) compared to the BAT arm (32.0%, Grade≥3 14.0%). The imbalance was mainly due to the differences among CMV infections.

However, in cGvHD, the higher proportion of patients with viral infections in the RUX arm (28.5%) vs the BAT arm (23.4%) up to Cycle 7 Day 1, was mainly due to the differences among BK virus infections.

Urinary tract infections (UTIs)

REACH-2 aGvHD Up to Day 28: were reported in 9.9% in the RUX arm and 10.7% in the BAT arm, and up to DCO in 19.1% vs 14.0%.

REACH-3 cGvHD Up to Cycle 7 Day 1: were reported in 8.5% in the RUX arm and 6.3 in the BAT arm, and up to DCO in 19.9 vs 9.5%.

Haematology/Cytopenias

aGvHD

REACH-2 population: At baseline, consistent with the study population disease characteristics, majority of patients with aGvHD had low haemoglobin level (90.8% vs. 90.6%), and low platelet counts (79.6% vs. 82.0%) in the RUX arm and the BAT arm; majority of these laboratory abnormalities were CTC grade 1 or 2. The worst post-baseline of grade 3 haemoglobin was reported in similar proportion of patients (47.4% vs. 46.0%). The proportion of patients with worst post-baseline value of grade 3 decreased platelets was lower in the RUX arm (25.7% vs. 32.7%) while higher for grade 4 decreased platelets (53.9% vs. 46.0%). Worst post-baseline of grade 3 (16.4% vs. 19.3%) and grade 4 (21.7% vs. 16.7%) neutrophil count was similar between the two arms.

Thrombocytopenia

Up to Day 28, 50.0% of the patients in the RUX arm and 32.7% of the patients in the BAT arm experienced at least one thrombocytopenia event. The incidence of Grade≥3 thrombocytopenia was 41.4% in the RUX arm and 29.3% in BAT arm, no fatal events were reported.

Up to the Secondary analysis data cut-off, the overall incidence of thrombocytopenia events increased to 56.6% (Grade≥3: 50.7%) in RUX arm and 36.7% (Grade≥3: 32.0%) in BAT arm. In 36.8% and 9.3%, in the RUX arm and the BAT arm, respectively the events were regarded as treatment related.

Anemia

Up to Day 28, the overall incidence of anemia events was 30.3% (Grade≥3: 22.4%) in the RUX arm and 28.0% (Grade≥3: 20.0%) in BAT arm.

Up to the Secondary analysis data cut-off, the incidence of anemia events had increased to 40.8% (Grade≥3: 36.2%) and 34.0% (Grade≥3: 25.3%). In 16.4% and 5.3%, in the RUX arm and the BAT arm, respectively, the events were regarded as treatment related.

Leukopenia

Up to Day 28, the overall incidence of leukopenia events was 32.9% (Grade≥3: 28.9%) in the RUX arm and 26.7% (Grade ≥3: 22.0%) in BAT arm.

Up to the Secondary analysis data cut-off, the incidence of leukopenia events had increased to 46.7% (Grade≥3: 42.2%) and 32.0% (Grade≥3: 27.3%).

cGvHD

REACH-3 population: At baseline, patients with cGvHD had low haemoglobin level (40.0% vs. 48.1%), and low platelet counts (39.4% vs. 34.9%) in the RUX arm and the BAT arm; majority of these laboratory abnormalities were CTC grade 1 or 2. Post-baseline values of CTC grade 3 haemoglobin decrease (17.6 vs. 9.5%) occurred more in the RUX arm than in the BAT arm. While post-baseline grade 3 neutrophil decrease (10.3% vs. 3.8%) was more frequent in RUX arm, grade 4 values were similar between the two arms (7.3% vs. 7.6%). Worst post-baseline platelet counts of grade 3 and grade 4was similar between the two arms [Study D2301 PA-Table 14.3-3.4]. More patients in the RUX arm had worst post-baseline grade 3 for haemoglobin decrease and neutrophil decrease.

Anemia

Up to Cycle 7 Day 1, the overall incidence of anemia events was 29.7% (Grade ≥3: 12.7%) in the RUX arm and 12.7% (Grade ≥3: 7.6%) in BAT arm.

Up to DCO, the incidence of anemia events was 32.1% (Grade ≥3: 15.2%) in the RUX arm and 13.9% (Grade ≥3: 7.6%) in the BAT arm.

Thrombocytopenia

Up to Cycle 7 Day 1, 21.2% of the patients in the RUX arm and 14.6% of the patients in the BAT arm experienced at least one thrombocytopenia event. The incidence of Grade ≥3 thrombocytopenia was 15.2% in the RUX arm and 10.1% in BAT arm.

Up to DCO, the overall incidence of thrombocytopenia was 23.0% (Grade ≥3: 17.6%) in RUX arm and 15.8% (Grade ≥3: 11.4%) in BAT arm.

Leukopenia

Up to Cycle 7 Day 1, the overall incidence of leukopenia events was 18.8 (Grade ≥3: 12.7%) in the RUX arm and 13.9 (Grade ≥3: 10.8%) in BAT arm.

Up to DCO, the incidence of leukopenia events was 23.6% (Grade ≥3: 18.8%) in the RUX arm and 14.6% (Grade ≥3: 11.4%) in the BAT arm.

Bleeding

aGvHD REACH-2

Up to Day 28, bleeding events were seen in 25.0% (Grade ≥3: 4.6%) and 22.0% (Grade ≥3: 2.6%) patients in the RUX arm and the BAT arm, respectively. There were no fatal events.

Up to the Secondary analysis data cut-off, the incidence of bleeding events was 40.1% (Grade ≥3: 12.5%) in the RUX arm and 28.0% (Grade ≥3: 7.3%) in BAT arm.

Corresponding exposure-adjusted IR of bleeding events were 158.2 and 174.3 per 100 PTYs in the RUX and the BAT arms respectively. One patient in ruxolitinib arm died due to lower GI hemorrhage in setting of aGvHD involving lower GI (stage 3). Events in 11.8% patients in ruxolitinib arm and 2.0% patients in BAT arm were regarded as treatment related. SAEs were observed in 6.6% patients in ruxolitinib and 5.3% patients in BAT arms. GI bleeding events occurred in 9.2% patients in the RUX arm and 6.7% patients in the BAT arm up. There were no fatal GI bleeding events. There was one intracranial haemorrhage reported, in each arm.

cGvHD REACH-3

Up to Cycle 7 Day 1, haemorrhages were seen in 11.5% (Grade ≥3: 2.4%) vs 14.6% (Grade ≥3: 1.9%) in the RUX arm and the BAT arm, respectively.

Up to DCO, hemorrhage events were seen in 12.7% (Grade ≥3: 3.0%) vs 16.5% (Grade ≥3: 3.2%) in the RUX arm and the BAT arm, respectively. There was one intracranial hemorrhage reported, in the RUX arm.

More bleeding events occurred in both treatment arms in REACH-2 than reported for REACH-3, concurrent with time elapsed since allo-tx and the higher frequency of thrombocytopenia.

Hypertension

REACH-2

Up to Day 28, hypertension occurred in 10.5% patients in the RUX arm and 9.3% patients in the BAT arm. Grade 3 events were observed in 5.9% patients in the RUX arm and 4.0% patients in the BAT arm, with no Grade 4 or 5 events.

Up to the Secondary analysis data cut-off, the proportion of patients with hypertension increased to 13.8% in the RUX arm, and 12.7% in BAT arm.

REACH-3

Up to Cycle 7 Day 1, hypertension occurred in 16.4% (Grade ≥3: 6.1%) and in 12.7% (Grade ≥3: 7.0%) RUX and BAT arms respectively.

Up to DCO, hypertension occurred in 18.2% (Grade ≥3: 6.7%) and in 13.3% (Grade ≥3: 6.7%) RUX and BAT arms respectively.

Hypertension rate increased by time on ruxolitinib and is an identified ADR for ruxolitinib.

Malignancies, including non-melanoma skin cancer

REACH-2: Up to the Secondary analysis data cut-off, malignancy events occurred in 3 (2.0%) patients each in two treatment arms. Corresponding exposure-adjusted IR of malignancy events were 6.1 and 9.9 per 100 PTYs in RUX and BAT arms respectively.

The event in RUX arm was diffuse large B-cell lymphoma (grade 3). On review of this case, the patient had relapse of underlying diffuse large B-cell lymphoma. In the BAT arm, one patient each had recurrent AML (fatal), erythroleukaemia (grade 1), and recurrent leukaemia (grade 4), none of which were suspected to be related to study treatment. After crossover, one patient had grade SAE of central nervous system lymphoma which led to discontinuation of study treatment.

REACH-3: Malignancy events were reported in 2 patients in the RUX arm and was not considered related to study treatment. These malignancy events were non melanoma skin cancer (NMSC) events. The event of grade 2 squamous cell carcinoma of skin was still ongoing by data cut-off in the Main treatment period. The events were reported as not related. No dose change, interruption or discontinuation occurred. The event of grade 3 basal cell carcinoma was resolved by Cycle 7 Day 1. No events were reported in the BAT arm.

Thromboembolic events

REACH-2: Up to Day 28, thromboembolic events occurred in 8.6% (Grade ≥3: 4.6%) and 6.7% (Grade ≥3: 4.7%) in RUX and BAT arms respectively.

Up to the Secondary analysis data cut-off thromboembolic events occurred in 12.5% (Grade ≥3: 7.2%) and 11.3% (Grade ≥3: 6.0%) in RUX and BAT arms respectively.

REACH-3: Up to Cycle 7 Day 1, thromboembolic events occurred in 4.8% (Grade ≥3: 2.4%) and 7.6% (Grade ≥3: 6.3%) in RUX and BAT arms respectively.

Up to DCO, thromboembolic events occurred in 6.1 (Grade ⩾3: 3.0%) and 7.6% (Grade ⩾3: 6.3%) in RUX and BAT arms respectively.

Clinical chemistry

Liver chemistry/Elevated Transaminases

REACH-2

Elevation of ALT, AST, bilirubin and alkaline phosphatase was noted in many patients. Of note, the liver is one of the target organs of underlying aGvHD indication. In majority of patients, the worsening post-baseline was to grade 1 or 2. a total of 8 (5.3%) in ruxolitinib arm and 18 (12.0%) patients in BAT arm met the criteria for Hy's law (concurrent ALT or AST >3x ULN and BILI >2x ULN and ALP <2x ULN). Confounders, like concomitant potentially hepatotoxic antibiotics or progressive liver GvHD, were found.

The following incidence data were reported for AEs by SOC in REACH-2 CSR:

Up to the DCO, 17.8% (Grade≥3: 6.6%) patients in the RUX arm and 10.0% (Grade≥3: 4.7%) patients in BAT arm had elevated transaminases. Events in 4.6% patients in ruxolitinib arm and 0.7% patients in BAT arm were suspected to be related to study treatment. There were dose reductions in

3.3% and dose interruption in 0.7% of the patients in the RUX arm. Drug-induced liver injury (DILI) was reported as AE in 2 patients in ruxolitinib arm and 1 patient in BAT arm and all were assessed as non-serious events with Grade 2 severity.

REACH-3

No liver enzyme abnormalities that met Hy's law were observed in the RUX arm. 4 patients in the BAT arm met the criteria for Hy's law. Out of these 4 patients, 3 of whom had baseline GvHD liver involvement and had significant elevation of liver tests indicating the elevation was pre-existing.

The following incidence data were reported for AEs by SOC in REACH-3 CSR:

Up to Cycle 7 Day 1 elevated transaminases were seen in 18.2% (Grade ≥3: 6.7%) of the patients in the RUX arm and 7.6% (Grade ≥3: 0.6%) in the BAT arm.

Up to DCO, elevated transaminases were seen in 20.6% (Grade ≥3: 8.5%) of the patients in the RUX arm and 7.6% (Grade ≥3: 0.6%) in the BAT arm.

Lipid abnormalities

REACH-2

Up to Day 28, lipid abnormalities were observed in 3.9% patients in ruxolitinib arm and 4.0% patients in BAT arm. Grade 3 abnormalities were observed in 1.3% patients each in both treatment arms. Treatment-related events occurred in 1.3% patients in the RUX arm only. One patient in BAT arm had a dose reduction. A total of 2.0% patients in each arm required medication in ruxolitinib arm and BAT arm.

Up to the Secondary analysis data cut-off, lipid abnormalities were observed in 9.9% patients in ruxolitinib arm and 7.3% patients in BAT arm.

REACH-3

Up to Cycle 7 Day 1, lipid abnormalities were observed in 18.8% (Grade ≥3: 7.3%) of the patients in the RUX arm and 14.6% (Grade ≥3: 7.0%) in the BAT arm.

Up to DCO, lipid abnormalities were observed in 20.6 (Grade ≥3: 7.3%) of the patients in the RUX arm and 14.6% (Grade ≥3: 7.0%) in the BAT arm.

Lipid abnormalities were overall higher in the RUX arm compared to the BAT arm. Increased total cholesterol, high-density lipoprotein (HDL) cholesterol, LDL cholesterol and triglycerides are known adverse drug reactions (ADRs) with ruxolitinib therapy.

Additional safety topics

Subgroup analyses for safety were conducted for aGvHD and cGvHD. The subgroups were by age, by gender, by race, by region, by baseline hepatic impairment status, by baseline renal impairment status, by concomitant cyclosporine treatment.

Safety assessment by age

REACH-2

Adolescents: There were 9 patients between 12 to <18 years (5 in the RUX arm and 4 in BAT arm). The median duration of exposure to RUX remained longer (163 days; range: 11.0 to 242.0) than to BAT (58.0 days; range: 2.0 to 162.0) and in line with overall population. The AE profile of adolescent

patients were similar to other age groups with minor differences. Cardiac disorders, vascular disorders, and immune system disorders were not observed in patients between 12 to 18 years treated with RUX. The profile of suspected AEs and SAEs were similar to other age groups. Among patients of 12 to <18 years, one patient died due to progression of underlying hematological disease. No fractures were reported in the RUX arm. Three patients reported events of bone pain, musculoskeletal pain (grade 1 events) and osteoporosis (n=1 each) (grade 2). Osteoporosis reported on study day 106 was resolving after treatment and did not result in a change in RUX therapy due to the event. All these events were assessed as not related to RUX therapy.

All ages: Due to the difference in the number of patients across age groups, 9 patients of 12 to <18 years (5 in the RUX arm and 4 in BAT arm), 248 patients of 18 to 65 years (126 in the RUX arm and 122 in BAT arm), and 45 patients of >65 years (21 in in the RUX arm and 24 in BAT arm) direct comparisons cannot be done.

In the 18 to 65 years old, there were 47 deaths (28.3%) and in the >65 years old group there were 14 deaths (48.3%).

REACH-3

Adolescents: There were only 11 patients in the age group 12-18 evaluable for safety, 4 received RUX and 7 BAT arm, 1 patient randomised to the BAT arm never received BAT dosing. Up to Cycle 7, Day 1, the median duration of exposure to RUX (25.6 weeks, range: 25.6 to 25.6) and to BAT (24.0 weeks, range: 8.9 to 25.6) was balanced. AEs profiles by SOC and PT were similar as observed in all-age group. No treatment-related AEs, and deaths and no fractures or bone abnormalities were reported. Up to Cycle 7 Day 1, 3 (75%) patients in the RUX arm and 7 (100%) patients in the BAT arm experienced at least one AE, of which, 1 patient in RUX arm and 2 patients in BAT arm experienced grade ≥3 AEs. AEs profiles by SOC and PT were similar as observed in all age group. Nervous system disorders (in 3 patients) and blood and lymphatic system disorders (2 patients) only occurred in the BAT arm. Immune disorders were not observed in patients between 12 to <18 years of age treated with RUX. SAE was reported in only one patient in the BAT arm with a nervous system disorders of grade 3 neuralgia and grade 2 transient ischemic attack. No death was reported for adolescent patients. No notable findings were observed in Tanner staging assessment for adolescent patients. Bone density measurement by DEXA scan were collected from adolescent patients. Results were not reported in this CSR. SAE was reported in only one patient No death was reported for adolescent patients.

All ages: Subgroup by age 18 to 65 years included 277 patients and subgroup by age ≥65 years included 40 patients. The AE profiles were similar between age groups with a few exceptions. A higher proportion of deaths were reported in the subgroup >65 year of age and mostly due to study indication (cGvHD and complications attributed to the treatment of cGvHD).

Safety analysis by organ involvement at randomization

The CSR for aGvHD summarizes that overall the profile of AEs was similar between the aGvHD organ involvements, except for the following differences: AEs by gastrointestinal disorders were more frequent in aGvHD patients with upper GI involvement and AEs by general disorders and administration site conditions were less frequent in aGvHD patients with liver involvement.

With respect to organ involvement at randomization in cGvHD, the overall population regardless of organs involved, the incidence of treatment-related AEs (any grade and grade ≥ 3), treatment related SAEs, AEs requiring dose adjustment/interruption, and AEs leading to discontinuation was higher in the ruxolitinib arm than in the BAT arm among patients with specific cGvHD organ involvement. However, this can be explained by the almost twice as high median total treatment exposure for the RUX arm.

Overall, when adjusted for exposure, the incidence of SAEs (all grades) was higher in the BAT arm (76.6 per 100 PTY) than in the RUX arm (60.3 per 100 PTY).

With respect to organ involvement at randomization, the profile of AEs was in general consistent with that of the overall population regardless of organs involved.

Vital signs

aGvHD: There were few notable changes in blood pressure or occurrence of increased pulse rate values are infrequent and comparable between treatment arms. Also changes in the weight were comparable between treatment arms. Decrease in weight was more frequent than weight gain in both treatment arms in aGvHD.

cGvHD: Changes in systolic blood pressure values was rare and comparable between treatment arms. However, the diastolic blood pressure changes were slightly different between 2 treatment arms: ≥105 mmHg and increase ≥15 mmHg occurred more frequently in ruxolitinib arm (4.8%, n=8) than in BAT arm (3.2%, n=5); ≤50 mmHg and decrease ≥15 mmHg occurred less frequently in ruxolitinib arm (3.0%, n=5) than in BAT arm (5.7%, n=9). This difference was enhanced in the Main treatment period up to the data cut-off.

The occurrence of increased pulse rate and corresponding AEs was comparable between two treatment arms and occurred in a low rate. Weight increase >10% was more common in the RUX arm (22.4%) compared to the in the BAT arm up to Cycle 7 Day 1 and even more pronounced up to DCO in the Main treatment period, 30.3% in the RUX arm and 16.5% in BAT arm.

Increase in weight was more frequent than weight loss in both treatment arms in contrast to the aGvHD population.

Exposure and AEs by age

REACH-2: There were 166 patients 18-65-year-old vs N=29 patients >65 years old. The median duration of exposure of RUX was lower in >65-year-old patients (4.3 weeks; range: 0.3-25.1) than in the 18 to 65 year-old group (9.9 weeks; range: 0.9 to 66.1). The corresponding exposure in PTY was 4.0 and 46.0 PTY. AEs by PT in all the subgroups were similar (18 to 65 years old group and >65 years old group) with a few exceptions. In the 18 to 65 years old group, cardiac failure and tachycardia were reported in 4 patients, and sinus tachycardia was reported in 3 patients. In the >65 years old group cardiac failure, tachycardia and sinus tachycardia were reported in one patient each. Occurrence of grade ≥ 3 AEs and occurrence of grade ≥3 suspected to study treatment were similar across the subgroups (18 to 65 years old group and >65 years old group). Patients with SAEs were higher the >65 years old group (82.8%) compared to the 18 to 65 years old group (66.9%). This was mainly due to infections and infestations (51.7%) (sepsis in 20.7% and pneumonia in 13.8%). Occurrence of all other SOCs and PTs were similar in the 18 to 65 years old group and in 65 years old group. In the 18 to 65 years old, there were 47 deaths (28.3%) and in the >65 years old group there were 14 deaths (48.3%). Among the 47 deaths in the 18 to 65 years old, 22 (13.3%) were due to study indication and 25 (15.1%) were due to other reason. The applicant has not provided this comparison for the BAT arm.

For cGvHD REACH-3, the median duration of exposure of RUX was similar across the subgroups (41.4 weeks in the 18 to 65 years old group (N=193); 42.6 weeks in the > 65 years old group (N=25). The AE profile was similar between age groups (12 to <18 years old group (N=8), 18 to 65 years old group and >65 years old group with few exceptions. Treatment related AEs (all grades and grade ≥3) were lower in 12 to <18 years old group compared to other subgroups. The proportion of patients with at

least one AE was 94.8% in 18 to 65 years old group, and 100% in >65 years old group. The proportion of patients with grade ≥ 3 AEs or SAEs was similar across the age groups. In study D2301, there were 15 deaths (7.8%) in the 18 to 65 years old group and 4 deaths (16.0%) in the 65 years old group. The applicant has not provided this comparison for the BAT arm.

Exposure and AEs by gender

REACH-2: the exposure duration in males (N=116) was 30.7 patient years and in females (N=85) was 21.1 patient. The AE profile (Occurrence of AEs related to study treatment, grade ≥3 AEs, SAEs, treatment related SAEs, fatal SAEs, AEs leading to discontinuation, AEs requiring dose adjustment/interruption, and AEs requiring additional therapy) was similar between the 2 genders. In REACH-1, the exposure in males (N=35) was 15.7 patient years and in females (N=36) was 9.9 patient years. The AE profile was similar between the 2 genders.

REACH-3: In Study D2301, the exposure duration in females (N=80) was 69.8 patient years and in males (N=146) was 125.3 patient years. The AE profile (Occurrence of AEs related to study treatment, grade ≥3 AEs, SAEs, treatment related SAEs, fatal SAEs, AEs leading to discontinuation, AEs requiring dose adjustment/interruption, and AEs requiring additional therapy) was similar between the 2 genders.

Exposure and AEs by race

REACH-2: The exposure duration was 35.4 patient years in White (N=141), 8.7 patient years in Asian (N=26), and 7.7 patient years in Others (N=34). In Study C2301, AEs and treatment related AEs were similar between Whites and Asian and lower compared to Others (56.8%). Grade ≥3 AEs were higher in Whites (93.7%) and Asian (92.9%) compared to Others (86.5%), SAEs were higher in Whites (74.9%) and Others (70.3%) compared to Asians (60.7%). AEs leading to discontinuation were generally comparable in all three races.

REACH-3: The exposure duration was 139.4 patient years in White (N=170), 35 patient years in Asian (N=36), and 20.7 patient years in Others (N=20). AEs were slightly higher in Others (100%) and Asian (97.2%) compared to Whites (93.5%). treatment related AEs were higher in Asian (77.8%) compared to Whites (56.5%) and Others (60.0%), grade ≥3 AEs were higher in Others (85%) compared to Whites (57.6%) and Asian (66.7%), SAEs were higher in Asian (55.6%) and Others (55.0%) compared to Whites (37.1%) AEs leading to discontinuation: higher in Asian (30.6%) compared to Whites (15.3%) and others (20.0%).

Exposure and AEs by region

REACH-2: The exposure duration was 4.6 patient years in Asia (excluding Japan), 3.7 patient years in Japan, and 43.5 patient years in Rest of the World. The number of patients in Rest of the World was highest (176) compared to patients in Asia (10) and Japan (15). Due to the differences in patient numbers, no conclusions could be made on the AEs by region.

REACH-3: The exposure duration was 10.3 patient years in Asia region, 21.3 patient years in Japan and 163.6 patient years in Rest of the World. The number of patients in Rest of the World was higher (194) compared to patients in Asia (8) and Japan (24). Due to the differences in patient numbers, no conclusions could be made on the AEs by region.

By baseline hepatic impairment status

REACH-2: In aGvHD, in the RUX arm (including cross over), 130 patients had no hepatic impairment, 21 patients with mild hepatic impairment, 27 patients with moderate hepatic impairment, and 23 severe hepatic impairment. Proportion of patients with SAEs in the no hepatic impairment group was 67.7% (n=88) and comparable between mild (n=16, 76.2%), moderate (n=17, 63.0%), and severe (n=18, 78.3%) groups. The fatal SAEs were 39 cases (21.9%) in the no hepatic impairment group, and 7 cases (28.0%) in the mild group, 14 cases (37.0%) in the moderate group, and 9 cases (39.1%) in the severe group. AEs leading to discontinuation in the no hepatic impairment group was 30.3% (54 cases) whereas in the mild, moderate and severe groups, it was 20.0% (5 cases), 38.2% (13 cases), and 28.6% (10 cases) respectively. No new safety concerns identified when safety data including hematological abnormalities were evaluated between patients with hepatic impairment and with no hepatic impairment at baseline.

In REACH-1, no notable differences were observed in the AE profile, laboratory abnormalities and clinical chemistry in patients with mild, moderate and severe hepatic impairment at baseline.

REACH-3 In cGvHD, the RUX arm (including cross over), most of the patients (n=205) had no hepatic impairment, 14 patients with mild hepatic impairment, 4 patients with moderate hepatic impairment, and 3 severe hepatic impairment. The proportion of patients with SAEs in the no hepatic impairment group was in 41.5% and was 42.9% in the mild group. One out of 4 patients had SAEs in the moderate (25%) and 2 out of 3 patients in the severe group (66.7%). Fatal SAEs in no hepatic impairment group were noted in 6.3%. In the mild and severe groups fatal SAEs were in 21.4% and 33.3%. There were no cases of fatal SAEs in the moderate group. Proportion of patients with AEs leading to discontinuation in the no hepatic impairment was 18.0%. In mild and severe groups, it was noted in 14.3% and 66.7%, respectively. In the moderate group, none of the 4 patients had an AEs leading to discontinuation. As there are few patients in the moderate and severe groups, these results should be interpreted with caution. No new safety concerns were identified.

By baseline renal impairment status

REACH-2: In aGvHD, 137 patients had no renal impairment at baseline, 49 in mild, 15 in moderate renal impairment groups. There were no patients with severe renal impairment. The proportion of patients with SAEs were 68.6% in no renal impairment group, similar to mild (69.4%) and moderate (73.3%) groups. The proportion of patients with fatal AEs were also similar across no renal impairment group (24.1%), mild (22.4%), and moderate (26.7%) groups. The AEs leading to discontinuation were in 28.5% in no renal impairment group, 34.7% in the mild group and in 20.0% in the moderate group.

REACH-3: In cGvHD, 117 patients had no renal impairment at baseline, 71 in mild, 35 moderate, and 2 severe renal impairment. The proportion of patients with SAEs were similar across the groups (41.9% in no renal impairment group, 38.0% in the mild group, and 42.9% in the moderate group), Also, the proportion of patients with fatal AEs were similar across the groups (8.5% in no renal impairment group, 7.0% in the mild group, and 5.7% in the moderate group). The AEs leading to discontinuation in the no renal impairment was 16.2%, similar to the moderate group (17.1%). In the mild group, AEs leading to discontinuation (22.5%) was slightly higher than in the no renal impairment.

Safety in cross-over population

It is acknowledged that the cross over groups (n=49) are expected to have more advanced disease. There were no new safety concerns identified between RUX randomized arm and crossover group in either pivotal study. In REACH 2, a similar proportion of patients had at least one AE in the ruxolitinib

arm (99.3%, grade ≥ 3: 91.4%) and in the cross-over group (95.9%, grade ≥3: 89.8%). The incidence of AEs, treatment related SAEs, and AEs requiring dose adjustment/interruption and additional therapy was also similar in the ruxolitinib arm and cross-over group. The incidence of SAEs, fatal SAEs and AEs leading to discontinuation were higher in the cross-over group than in the ruxolitinib randomized group but notably treatment related events were similar. In REACH 3 the incidence of SAEs, fatal SAEs, AEs leading to treatment discontinuation was lower to that of the ruxolitinib randomized arm. Thus overall, it can be concluded that in general the safety of the cross over group is similar for the cross over patients

Post marketing experience

Published information on the safety of RUX is evaluated on an ongoing basis and is available in the PSURs for RUX, with PSUR 10 covering the period 23-Feb-2019 to 22-Feb-2020, please refer to section 5.

RUX was approved on 16 Nov 2011 for treatment of MF. An estimate of patient exposure is calculated based on worldwide sales volume in kilograms (kg) active substance sold during the reporting interval and the Defined Daily Dose of 30 mg. The sales volume of RUX cumulatively up to 22 Feb 2020 (PSUR 10 data cut off) since the IBD of the product was estimated to be approximately 1,670.75 kg active substance (Novartis 1175.20 kg and Incyte 495.55 kg). The estimated cumulative exposure was approximately 152,580 PTY as discussed in PSUR 10 (reporting period 23-Feb-2019 to 22-Feb-2020). The Risk Management Plan (RMP) continues to adequately describe the measures to manage risk.

2.6.1. Discussion on clinical safety

Due to the differences in study populations in acute GvHD and chronic GvHD, the clinical manifestation and time elapsed since transplant, the extent of exposure is markedly different between the diagnoses. This accounts for the RUX arm but also for the BAT arms. Many AEs can be difficult to distinguish from symptoms of aGvHD or by the transplant procedure itself. The exposure differed significantly between the acute GvHD studies and the chronic GvHD setting. In aGvHD (REACH-2) the median duration of exposure to RUX was 63.0 days (range: 6.0 to 396.0) and for BAT 29.0 days (range: 1.0 to 188.0). In contrast, in the chronic setting the median duration of exposure was 42.6 weeks (range: 21.7-74.6) in the RUX arm and 25.2 (range: 22.7-41.1) weeks in the BAT arm. Due to the differences in the study populations, study structure in the Phase 3 studies and differences in AE frequency and patterns, the safety data are separately presented for aGvHD and cGvHD. Overall, the exposure to RUX in presented studies was considered sufficient to allow for a primary assessment of safety in patients representative of the intended target population, long-term safety data has not been presented and is listed as missing information in the RMP.

It is noted that out of the selected AESIs, growth retardation, PML and tuberculosis, there were no patients under these safety topics in any of the three presented studies. Other selected AESIs, e.g., dizziness, headache, flatulence, constipation and facture, were either not reported or low and comparable between the treatment arms.

The comparator in respective study, consist of 9 and 10 subsets of treatment choices for aGvHD and cGvHD, respectively, of whom some treatment choices are the same and some are unique for the respective study.

In REACH-2 the rate of discontinuation until day 28 was acceptable with 11.2% discontinuing due to an AE (mostly grade 3) with as main AE anaemia and thrombocytopenia. The majority of patients with aGVHD required at least one dose change or interruption of RUX (82.9%), with as primary reason AEs

(57.2%), additionally a large proportion of subjects had a dose interruption by physician decision (19.1%) thus not classified as a discontinuation due to AE or lack of efficacy. However, the CRFs did not collect the reason for the physician decision and as such the reason for dose change or dose interruption cannot be analysed for approximately 20% of the study population.

The side effects profile in the ruxolitinib arm is mainly characterised by hematological adverse events and infections, consistent with the mechanism of action of ruxolitinib and these are known ADRs for ruxolitinib. However, due to the new indications proposed, aGvHD and cGvHD, which are two populations much more vulnerable with respect to both cytopenias and infections, these ADRs are also the main ADRs in the comparator arm.

With respect to the short time, 28 days, to the primary endpoint in REACH-2 study compared to the time up to primary endpoint in the cGvHD study REACH-3, Cycle 7 Day 1, and in addition, for both studies the double median exposure time on ruxolitinib compared to the BAT arm, complicated by the variety of BAT choices, comparisons between the two studies and between treatment arms is challenging. Hence, incidence of AE up to the data cut-off should be viewed with consideration of duration of exposure, population studied as well as the spectrum of BAT treatments.

By time of Secondary analysis, DCO for aGvHD the most commonly affected AEs by SOC were infections and infestations (81.6% of patients) and blood and lymphatic system disorders (71.1% of patients) for the RUX arm and for the BAT arm, 71.3% and 51.3% for SOC infections and infestations and blood and lymphatic system disorders, respectively.

The pattern for cGvHD is similar, namely by time of DCO the Main treatment period for cGvHD the most commonly affected AEs by SOC were infections and infestations (71.5% of patients) and blood and lymphatic system disorders (46.1% of patients) for the RUX arm and for the BAT arm, 65.8% and 24.1% for SOC infections and infestations and blood and lymphatic system disorders, respectively. Analysis of AEs that were judged as permanent was based mainly on AEs that were not recovered/not resolved in patients who discontinued and who have finished their 30-day safety follow-up.

In both REACH-2 and RECAH-3, the most frequent AEs that were not resolved were cytopenias, all other SAEs were reported infrequent and with no specific pattern and not leading to new safety concerns.

Furthermore, there is a considerable difference with respect to SAEs, fatal SAEs, AEs leading to discontinuations and AEs leading to dose adjustment/interruption in relation to ruxolitinib treatments in acute and chronic GvHD.

Since the incidences of AEs by SOC and by PT differs, in some respects, from those for the already approved ruxolitinib indications, a separate ADR table presenting data for the two pivotal studies in GvHD and the frequencies (all grades and grade 3/4), is introduced in the SmPC. This is endorsed. Cross over from BAT towards RUX was allowed in the study and it is reassuring that the type of AEs for cross over patients was similar and there were no new safety concerns identified and therefore, inclusion in the table in SmPC 4.8 is acceptable.

The results for primary cause of death did not show any new safety signals compared to the primary analysis results. A detrimental effect on OS of ruxolitinib cannot be concluded based on still immature data (please refer to efficacy section).

The rate of malignancy relapse/progression was roughly similar for ruxolitinib and BAT in each of the pivotal trials.

Subgroups

Subgroup analyses of patients with baseline hepatic or renal impairment in REACH - and REACH-3 were provided. There was no clear trend that patients with more severe baseline hepatic or renal impairment are at increased risk, but the assessment is hampered by sample size and numbers were not robust over the categories. Importantly, no new safety concerns were identified in patients with severe hepatic or renal impairment and the starting RUX dose will not be reduced, as the dose is already low, and there is a need to ensure an adequate dose for efficacy. For renal impairment no specific dose adjustments are advised for mild or moderate renal impairment, which is accepted.

The comparison with BAT in the elderly age group shows that in REACH 2, patients older than 65 experience more grade 3 or higher AEs in RUX arm (95.2%) than in BAT (79.2%) and more were considered treatment related in the RUX arm. This was also confirmed with higher percentages SAE, AEs leading to discontinuation, AEs leading to dose adjustments and AEs requiring additional therapy. However, the sample size limits accurate definition of incidences.
A conclusion of ruxolitinib safety evaluation in adolescents is not readily feasible due to the low number of participants. With respect to patients >65 year of age, besides also a more limited number, these patients are already selected, with regard to few or any co-morbidities, before entering the allo-transplant procedure. However, no data points towards a significant different AE profile between the age groups. The death rate is higher in patients >65 years in both the aGvHD and cGvHD study.

Long term safety aspects

In study REACH-2, one patient experienced a DLBCL (grade 3) up to data cut-off which was assessed not related to RUX treatment. In the chronic GvHD population REACH3, up to Cycle 7 Day 1, malignancy events were reported in 1.2%, judged as not related to RUX and no new events reported at the time of the data cut-off. Thus, within this reporting period the risk for malignancies is low even with prolonged exposure (as in cGvHD). Long-term safety data including secondary malignancies are listed in the RMP as missing information and relevant information are incorporated in SmPC section 4.4. and 4.8.

Assessment of paediatric data on clinical safety

In study REACH-2, 3 patients reported events of bone pain, musculoskeletal pain and osteoporosis (n=1 each), with low grade and 1 patient reported osteoporosis as a grade 2 event (on study day 106) resolving after treatment and considered not related to RUX. In REACH-3, no notable findings were observed in Tanner staging assessment. The non-clinical data showed that ruxolitinib administration to juvenile rats (PND 21, roughly corresponding to 2-year-old children) was associated with reduced bone size (diameter and/or lengths). Serious adverse effects consisting of degeneration of the physis and physeal fractures have been observed at higher systemic exposures when starting to dose from PND 7 (corresponding to preterm/neonate). Results in adolescent patients with sequential DEXA scan results revealed no clinically relevant change in the bone mineral density during the study, although data was limited to 4 subjects in RUX arm and 4 subjects in BAT arm. Additionally, no clinically relevant change in the weight and height was noted, which is reassuring.

Thus, the available DEXA scans and data on growth and development in adolescents in REACH-2 and REACH-3 studies did not confirm the safety concerns associated with bone toxicity. Moreover, the clinical relevance of the non-clinical results for the current indication for adults and adolescents is limited. Considering the information available in paediatric patients (<12 years) is very limited at this point in time, the applicant adds this information as missing information in the RMP. Overall, the safety profile in adolescents could be considered similar to adults, although uncertainties are identified with respect to the small number of subjects and reassurance on the potential for bone toxicity is needed. When available Tanner and bone density measures and preliminary safety assessment of ongoing

paediatric studies (Study INC424F12201 [REACH-4] in aGvHD and Study INC424G12201 [REACH-5] in cGvHD) are requested.

2.6.2. Conclusions on clinical safety

The overall safety profile of RUX-treated patients in both acute and chronic is consistent with its established safety profile and as expected in the study population. The safety profile of Jakavi is mainly characterised by cytopenias and infections. No new safety concerns were identified in the present GvHD studies with ruxolitinib therapy; however, ruxolitinib may be less tolerable in the GvHD setting compared to that of myelofibrosis or PV but appears to be manageable with the dose modification guidance.

2.6.3. PSUR cycle

The requirements for submission of periodic safety update reports for this medicinal product are set out in the list of Union reference dates (EURD list) provided for under Article 107c(7) of Directive 2001/83/EC and any subsequent updates published on the European medicines web-portal.

2.7. Risk management plan

The MAH submitted an updated RMP version with this application.

The CHMP received the following PRAC Advice on the submitted Risk Management Plan:

The PRAC considered that the risk management plan version 14.0 is acceptable.

The CHMP endorsed this advice without changes.

The CHMP endorsed the Risk Management Plan version 14.0 with the following content:

Safety concerns

Summary of safety concerns

Important identified risks	Serious infections
Important potential risks	Developmental toxicity Non-melanoma skin cancer (including basal, squamous and Merkel cell carcinoma)
Missing information	Long-term safety data, including secondary malignancies Safety in pediatric patients ≥12 years (GvHD only)

Pharmacovigilance plan

Ongoing and planned additional pharmacovigilance activities

Study Status	Summary of objectives	Safety concerns addressed	Milestones	Due dates
Category 1 - Imposed mandatory additional pharmacovigilance activities which are conditions of the marketing authorization				
None				

Study Status	Summary of objectives	Safety concerns addressed	Milestones	Due dates
Category 2 – Imposed mandatory additional pharmacovigilance activities which are Specific Obligations in the context of a conditional marketing authorization or a marketing authorization under exceptional circumstances				
None				
Category 3 - Required additional pharmacovigilance activities				
Interventional study Study INC424F12201 Ongoing	Phase I primary objective: To assess PK parameters (i.e. AUC, Cmax, T1/2, Ctrough) of ruxolitinib for patients with aGvHD and SR-aGvHD. Phase II primary objective: To measure the activity of ruxolitinib in patients with grade II-IV aGvHD or grade II-IV SR-aGvHD assessed by ORR at Day 28.	Safety in pediatric patients ≥12 years (GvHD only).	Date of initiation Final CSR	Feb-2019 31-Mar-2024
Interventional study Study INC424G12201 Ongoing	To evaluate the activity of ruxolitinib added to standard dose corticosteroids, ± CNI, in pediatric subjects with moderate or severe treatment naive-cGvHD or SR-cGvHD	Safety in pediatric patients ≥12 years (GvHD only).	Date of initiation Final CSR	May-2020 31-Dec-2025

Risk minimisation measures

Summary of risk minimization activities by safety concerns

Safety concern	Risk minimization measures
Important identified risks	
Serious infections	Routine risk minimization measures: SmPC Section 4.4: Precaution for monitoring, treatment and description of risk factors and nature of risk. Section 4.8: The ADRs of UTI, HZ, pneumonia and sepsis are listed. Additional risk minimization measures: None.
Important potential risks	
Developmental toxicity	Routine risk minimization measures: SmPC Section 4.1 Section 4.2

Safety concern	Risk minimization measures
	Section 4.3 Section 4.6 Section 5.3 There are no data from the use of ruxolitinib in pregnant women. Additional risk minimization measures: None.
Non-melanoma skin cancer (including basal, squamous and Merkel cell carcinoma)	Routine risk minimization measures: SmPC Section 4.4 Precautionary instructions on monitoring of patients who are at increased risk of skin cancer. Additional risk minimization measures: None.
Missing information	
Long-term safety data, including secondary malignancies	Routine risk minimization measures: The safety profile of ruxolitinib is described in SmPC Section 4.8. Currently available data do not support the need for additional risk minimization. Additional risk minimization measures: None.
Safety in pediatric patients ≥12 years (GvHD only)	Routine risk minimization measures: SmPC Section 4.2: The ruxolitinib dose in pediatric patients with GvHD aged 12 years and older is the same as in adults. Additional risk minimization measures: None.

2.8. update of the Product Information

As a consequence of this new indication, sections 4.1, 4.2, 4.4, 4.8, 5.1 and 5.2 of the SmPC have been updated with regards to acute and chronic GvHD. The Package Leaflet has been updated accordingly. In addition, the list of local representatives in the PL has been revised to amend contact details for the representative of The Netherlands.

Please refer to Attachment 1 which includes all agreed changes to the Product Information.

2.8.1. User consultation

A justification for not performing a full user consultation with target patient groups on the package leaflet has been submitted by the MAH and has been found acceptable for the following reasons:

The information proposed in the Package leaflet maintain the currently approved layout and format and are not considered to require further consultation with target patient groups. The key information remain the same as for the currently approved PL or contain editorial changes only.

3. Benefit-Risk Balance

3.1. Therapeutic Context

3.1.1. Disease or condition

Allogeneic hematopoietic cell transplantation (HCT) is a potentially curative treatment option for a variety of hematologic malignancies and also several non-malignant hematologic diseases.

Acute GvHD develops in 50% to 70% of patients after alloSCT with conventional prophylaxis and is one of the major barriers to successful transplant outcomes. The median time to onset (grade II-IV) of 20-25 days after engraftment.

The pathogenesis of aGVHD is complex and is initiated when alloreactive donor immune cells recognize immunologically disparate antigens in the host. The risk of developing aGVHD depends on the degree of HLA match, recipient age, graft source, underlying disease diagnosis, intensity of conditioning regimen and GVHD prophylaxis used.

The clinical manifestations are seen primarily in three organs: the skin (maculopapular erythematous skin rash, erythroderma), the liver (cholestasis, hyperbilirubinemia, and/or jaundice), and the lower and upper gastrointestinal tract (nausea, abdominal pain, vomiting, anorexia with weight loss, secretory diarrhea, GI bleeding and/or ileus).

Patients who develop steroid-refractory acute graft-versus-host disease (SR-aGVHD) after allogeneic hematopoietic cell transplantation have poor prognosis. Systemic corticosteroids are the recommended first-line treatment of grades II to IV aGVHD, but less than 50% of patients achieve durable responses. The reported 6-month survival estimate for patients with steroid-refractory aGVHD is approximately 50%, with 30% or less of patients surviving beyond 2 years

Chronic GvHD is a major long-term complication after alloSCT, occurring most frequently after 100 days post-transplant with a median time to onset reported as 162 days post-transplant. While aGvHD is mainly a mature donor T cell-mediated inflammatory disease, cGvHD is characterized by the activation of complex signalling pathways in both T and B cells, reduced levels of circulating regulatory B cells (Bregs) and CD4+ Tregs

Among patients who undergo alloSCT, cGvHD occurs in 30% to 70% of patients. The occurrence of cGvHD varies depending on the donor type. Approximately 30% of cGvHD are de novo without any preceding aGvHD.

Chronic GvHD is a leading cause of non-relapse mortality and morbidity in patients surviving more than 2 years after transplantation and unfavourably affects physical and functional well-being as well as quality of life of most of the patients who are otherwise cured for their underlying disease after HSCT. Mild cGvHD is associated with a good prognosis whereas moderate and severe disease are associated with higher treatment-related mortality and lower survival. The 2-year overall survival (OS) reported for mild, moderate, and severe disease is 97%, 86%, and 62%, respectively.

Children are at less risk for GVHD than adults; however, that risk is still significant especially when using alternative donor sources. The incidence of grade II–IV aGvHD in children ranges from 28 to 56%, depending on the degree of histocompatibility, recipient age, underlying condition, and conditioning regimen used. The mortality in adolescent patients ≥ 12 years of age is, however, similar to that in adults (49% survive beyond 6 months). Grade II-IV SR-aGvHD adult

Chronic GvHD usually involves not only the epithelial target tissues affected in classic aGvHD (GI tract,

liver, skin,) but also additional organ systems including lungs, muscles, fascia, joints, genitalia, eyes, and nails.

The final approved indication is:
Jakavi is indicated for the treatment of patients aged 12 years and older with acute graft versus host disease or chronic graft versus host disease who have inadequate response to corticosteroids or other systemic therapies.

3.1.2. Available therapies and unmet medical need

There are no uniformly used, second-line treatments following corticosteroid resistance or sustained need, in SR-GvHD. Ibrutinib has been approved by FDA and Health Canada for the treatment of adult patients with cGvHD after failure of one or more lines of systemic therapy. Ruxolitinib is registered for the treatment of aGvHD in the US since 2019, and therefore, practices vary as to the selection of various systemic therapies but also depending on availability (e.g., ECP). There are no products approved for steroid refractory GVHD in the EU.

3.1.3. Main clinical studies

The Applicant provided two pivotal studies to support the two claimed indications,

- CINC424C2301 (REACH-2): A phase III randomized open-label multi-centre study of ruxolitinib versus BAT in patients with corticosteroid-refractory acute graft vs. host disease after allogeneic stem cell transplantation. Initiated 10-Mar-2017 and a DCO of 25-Jul-2019 for the primary analysis and 06-Jan-2020, DCO for second analysis. The study is ongoing.

- CINC424D2301 (REACH-3): A phase III randomized open-label multi-centre study of ruxolitinib versus BAT in patients with corticosteroid-refractory chronic graft vs. host disease after allogeneic stem cell transplantation. Study initiation date: 29-Jun-2017 and DCO 08-May-2020 for the primary analysis. The study is ongoing.

A supportive Phase 2 study, INCB 18424-271 (REACH-1) was performed in SR-aGvHD.

The choice of best available therapy (BAT) for the control group in the two presented pivotal studies, was decided by the Investigator, from a list defined in the protocol before randomization in respective study.

Some treatment options were the same for aGvHD and cGvHD (Extracorporeal photopheresis (ECP)low-dose methotrexate (MTX) mycophenolate mofetil (MMF) mTOR inhibitors (everolimus or sirolimus) infliximab, while the choices anti-thymocyte globulin (ATG), mesenchymal stromal cells (MSC) and etanercept were unique for REACH-2 (aGvHD); and rituximab, pentostatin, imatinib and ibrutinib were unique choices for REACH-3 (cGvHD).

3.2. Favourable effects

Pivotal study REACH-2; ruxolitinib versus vs BAT in SR-aGvHD

- ORR on Day 28 (primary endpoint, primary analysis) was 62.3% (95% CI: 54.2, 70.0) in the RUX arm and 39.4% (95% CI: 31.6, 47.5) in the BAT arm ($p<0.0001$, stratified Cochrane-Mantel-Haenszel test, one-sided, odds ratio: 2.64 with 95% CI: 1.65, 4.22). The results were maintained at the Second analysis.

Subgroup analysis (follow-up data up + 6 months) indicated consistent results with that of the Primary analysis results. The OR favoured ruxolitinib across most of baseline characteristics subgroups.

- Durable ORR at Day 56 (key secondary endpoint, primary analysis) were in favour of ruxolitinib (OR: 2.38; 95% CI: 1.43, 3.94; p=0.0005). Results were maintained at the second analysis.
- Median DOR was longer in RUX arm (163 days, range: 22.0 to 623.0) than in BAT arm (101 days, range: 10.0 to 456.0).
- At Day 56, more patients in ruxolitinib arm (22.1%) had tapered off corticosteroids than in BAT arm (14.8%)

Pivotal study REACH-3; ruxolitinib vs BAT in SR-cGvHD

- The ORR (Cycle 7 Day 1, primary endpoint, primary analysis) was 49.7% (95%CI: 41.8, 57.6) in the RUX arm and 25.6% (95%CI: 19.1, 33.0) in the BAT arm, odds ratio = 2.99 (95% CI: 1.86, 4.80) (stratified CMH test p<0.0001, one-sided).
- FFS (failure free survival) showed a hazard ratio = 0.370 (95% CI: 0.268, 0.510), p<0.0001 in favour of the RUX arm. The 6-months FFS probability was 74.89% (95% CI: 67.48, 80.85) in the RUX arm and 44.46% (95% CI: 36.46, 52.14) in the BAT arm.
- Significant improvement in TSS, cycle 7 day 1 (second Key secondary endpoint) was shown for the RUX arm compared to the BAT arm at the primary analysis. The odds ratio (RUX/BAT) was 2.62 (95% CI: 1.42, 4.82, stratified CMH test, p = 0.0011).
- Up to Cycle 7 Day 1, a numerically larger proportion of patients in the RUX arm (24.2%) had completely tapered off corticosteroids vs. 16.8% in the BAT arm.

In GvHD paediatric patients (12 years of age and older), the safety and efficacy of Jakavi are supported by evidence from the randomised phase 3 studies REACH2 and REACH3 (see section 4.2 for information on paediatric use). In REACH2, responses were observed at day 28 in 4/5 adolescent patients with acute GvHD (3 had CR and 1 had PR) in the ruxolitinib arm and in 3/4 adolescent patients (3 had CR) in the BAT arm. In REACH3, responses were observed at cycle 7 day 1 in 3/4 adolescent patients with chronic GvHD (all had PR) in the ruxolitinib arm and in 2/8 adolescent patients (both had PR) in the BAT arm.

3.3. Uncertainties and limitations about favourable effects

Both of the pivotal studies are characterised by high numbers of protocol deviations (PD), relating e.g., to aGVHD grading/staging, response assessment and the use of prohibited medication, which raised concerns, not least due to the open-label nature of the studies. For REACH-2, 36.4% of the participants in the RUX arm and 43% in the BAT and for REACH 3, 41.8% in the RUX arm and 43.9% in the BAT arm, were excluded from the sensitivity analyses. A GCP inspection found that the process for PD management was not robust enough and should have been improved in order to enhance a timely resolution of PD, e.g., related to prohibited medication and organ staging assessment done per investigator criteria/judgement rather Harris.

Furthermore, with regard to BAT treatment accountability, which was not mandatory per study protocols, the inspection found that the Applicant failed to assure for non-IMP appropriate traceability and compliance (i.e., not compliant with the Guidance documents applying to clinical trials guidance on investigational medicinal products (IMPs) and 'non investigational medicinal products' (NIMPs) (rev. 1,

March 2011)"). In the worst scenario for the REACH 2, 12% of the trial participants and for the REACH 3, 41% of the trial participants receiving BAT, treatment compliance was not/insufficiently monitored. This introduces some uncertainty about estimates.

aGVHD

For the primary endpoint of REACH-2, patients were assigned as failing to reach OR, if there was a need to intensify immunosuppression. Thus, it is presently unclear whether to understand the contribution of Jakavi to the treatment armamentarium is mainly as an alternative treatment option, and to what extent it contributes to a greater total treatment efficacy

Most responses noted at ORR Day 28, had already occurred at Day 14 (other secondary endpoint). However, ORR Day 56 (Durable ORR at Day 56) reveal that ORR dropped by 37% in the RUX arm and by 44% in the BAT arm between these two datum points, i.e. early responses, but for at large subset of responding patients, a short duration. Thus, there is some uncertainty about the extent of benefit shown.

cGVHD

Even though an updated OS analysis has been performed, data are still immature,

For the primary endpoint, patients were assigned as failing to reach OR, if there was a need to intensify immunosuppression. Thus, it is unclear whether to understand the contribution of Jakavi to the treatment armamentarium is mainly as an alternative treatment option or to what extent it contributes to a greater total treatment efficacy.

FFS is a composite endpoint, as described above. As anticipated, this endpoint is driven by impact on addition or initiation of another systemic therapy, while there is no indication of any impact of treatment on relapse of underlying disease or non-relapse mortality.

For REACH 3 most responses noted for the primary endpoint ORR Cycle 7 Day 1, had already occurred at Cycle 4 Day 1 (other secondary endpoint). The endpoint BOR shows that one third of the initial responses to ruxolitinib were no longer recorded at the timepoint of the primary endpoint in cGvHD. However, early discontinuation, deaths, changes in systemic cGvHD treatment before Cycle 7 Day 1 or missing visits were the main reasons for the differences in BOR and ORR therefore, not necessarily meaning treatment failure or loss of response in all cases.

The patient reported outcome TSS (second key secondary endpoint) derives from an open-label study. The risk of bias due to this cannot be measured or ascertained. Furthermore, the "mLSS" symptom scale on which TSS was based, according to the GCP inspectors, was not validated prior to start of the trial.

REACH-2 and REACH-3

Efficacy in patients with low grade GvHD and in patients who received more than one prior systemic treatment for SR GvHD has not been investigated (for both pivotal trials).

The number of investigated adolescent patients is very limited (n=9 in total for both pivotal trials), emphasizing the need for extrapolation of adult data.

Response rates seem similar to BAT or even more in favour of BAT in small subgroups of patients >65 years of age (aGvHD), region Asia (a+cGvHD) and some subgroups with prior steroid + CNI + other systemic therapy for a or cGvHD. With regard to this last subgroup, heterogeneity in previously

allowed as well as concomitant treatment and prophylaxis regimes for GvHD make it difficult to interpret results.

The impact of RUX in terms of OS is uncertain since interpretation of data is hampered by cross-over in >30% of patients in both trials.

Efficacy in patients with overlap syndrome, GvHD after pre-emptive treatment of malignancy recurrence with donor lymphocyte infusion and patients that did not tolerate steroid treatment is unknown.

3.4. Unfavourable effects

aGvHD REACH-2

The side effects profile by SOC is mainly characterised by hematological adverse events and infections. Other AEs by SOC frequently seen in were investigations, GI-disorders, general disorders, metabolism and nutrition and respiratory, thoracic and mediastinal disorders.

- AEs leading to treatment discontinuations Up to Day 28: 11.2% and 4.0% in the RUX arm and BAT arm, respectively.

  Up to the second DCO, AEs leading to discontinuation were observed in 27% in the RUX arm and 9.3% in the BAT arm.

- AEs leading to Dose adjustment or interruption up to Day 28, 36.8% and 9.3% the RUX arm and BAT arm, respectively.

  Up to DCO AEs (all grades) leading to dose adjustment or interruption, was 54.6% in the RUX arm and 13.3% in the BAT arm Most of the AEs were of grade ≥3 in severity (in (50.0% and 9.3%), driven by cytopenias.

- Deaths up to Day 28: there were 9.9% and 14.0% on-treatment deaths recorded in RUX arm and the BAT arm, respectively. The primary reason for death was aGvHD.

  Up to DCO there were a total of 168 deaths, 53.9% deaths in the RUX arm and 57.3% deaths in the BAT arm. Deaths due to aGvHD occurred in 24.3% in the RUX arm and 25.3% in the BAT arm.

- SAEs Up to Day 28: 37.5% in the RUX arm and 34.0% in the BAT arm. By SOC: Infections and infestations (21.7% vs. 17.3%) and by PT, Sepsis (5.3% vs. 2.0%), CMV infection (2.6% vs. 3.3%), Respiratory failure and septic shock (2.6% vs. 2.7%, each) and Pneumonia (2.0% each)

  SAEs Up to DCO were reported in 66.4% patients in RUX arm and 53.3% in BAT arm. The difference was mainly driven by the SOC of infections and infestations (38.2% in the RUX arm and 30.0% in the BAT arm)

AESIs

- Hematological AEs up to Day 28, with respect to hematological events, were recorded in 58% (Gr ≥3: 47.4%) in the RUX arm and 44.7% (Gr ≥3: 4.7%) in the BAT arm.

  AEs by PT, up to Day 28, were thrombocytopenia (50.0% vs. 32.7%) anemia (30.3% vs 28.0%), leukopenia (32.9% vs 26.7%) neutropenia (15.8% vs. 12.7%) and neutrophil count decreased (6.6% vs 10.0%) in the RUX arm and the BAT arm, respectively.

Up to the Secondary analysis DCO, the overall incidence of thrombocytopenia events increased to 56.6% (Grade≥3: 50.7%) in RUX arm and 36.7% (Grade≥3: 32.0%) in BAT arm the overall incidence of anemia 40.8% (Grade≥3: 36.2%) in the RUX arm and 34.0% (Grade≥3: 25.3%) in the BAT arm leukopenia in 46.7% (Grade≥3: 42.2%) and 32.0% (Grade≥3: 27.3%) in the RUX arm and the BAT arm, respectively.

- AEs Infections Up to Day 28, were reported in 61.2% the RUX arm vs 58.7% the BAT arm. The proportion of opportunistic infections during period up to Day 28 was 27% vs 22% in the RUX vs BAT arm and cytomegalovirus infection (25.7% vs. 20.7%).

  AE pneumonia up to Day 28 was seen in 27% vs 22% in the RUX vs BAT arm and CMV infections were reported in 28.3% in the RUX arm vs. 24.0% the BAT arm.

cGvHD REACH-3

Similar to aGvHD as well as the other studied indications, the side effects profile is mainly characterised by hematological adverse events and infections. Other AEs by SOC frequently seen in were investigations, GI-disorders, general disorders, metabolism and nutrition and respiratory, thoracic and mediastinal disorders as in the pivotal for aGvHD.

- AEs leading to Treatment discontinuations: Up to Cycle 7 Day 1: AEs leading to study treatment discontinuation were 16.4% vs. 7.0% in the BAT arm. Pneumonia was the most common (≥ 2% patients) AE leading to discontinuation in both arms.

  Up to DCO, AEs (all grades) leading to discontinuations were reported in 18.1% and grade ≥3 AEs in 13.3% of ruxolitinib treated patients.

- AEs leading to Dose adjustment or interruption up to Cycle 7 Day 1: AEs leading to dose adjustment or interruption were 37.6% vs. 16.5% in the RUX arm and the BAT arm respectively.

  AEs leading to Dose adjustment or interruption Up to DCO for the main treatment period occurred in 43.0% in the RUX arm and 18.4% in the BAT arm.

- Deaths Up to Cycle 7 Day 1: 7.9% patients in the RUX arm and 5.7% patients in BAT arm died on treatment. Study indication, cGvHD, was the main cause of death.

  Deaths Up to the DCO were 18.8% in the RUX arm and 16.5% in the BAT arm.

- SAEs Up to Cycle 7 Day 1, were observed in 33.3% in the RUX arm and 36.7% in the BAT arm. By PT, the most frequent SAEs were pneumonia, pyrexia, lower respiratory tract infection, and bronchopulmonary aspergillosis.

  SAEs Up to DCO were seen in 43.6% of patients in the RUX arm and 39.9% in the BAT arm.

AESIs

- Hematological AEs up to Cycle 7, Day 1, were recorded in 41.8% (Gr ≥3 23.0%) in the RUX arm and 22.2% (Gr ≥3 15.8%).

  AEs by PT, up Cycle 7 Day 1, subjects with at least one event of thrombocytopenia were 21.2% vs. 14.6%, anemia 29.7% vs 12.7%, leukopenia 18.8% vs 13.9%) in the RUX arm and the BAT arm, respectively.

  Up to the Secondary analysis DCO, events of thrombocytopenia were 23.0% (Grade ≥3: 17.6%) in the RUX arm and 15.8% (Grade ≥3: 11.4%) in the BAT arm, the overall incidence of anemia was 32.1% (Grade ≥3: 15.2%) in the RUX arm and 13.9% (Grade ≥3: 7.6%) in the

BAT arm, leukopenia in 23.6% (Grade ≥3: 18.8%) and 14.6% (Grade ≥3: 11.4%) in the RUX arm and the BAT arm, respectively.

- AEs Infection Up to Cycle 7 Day 1, were reported for 62.4% vs. 58.2% in the RUX arm and the BAT arm respectively.

  The most frequently reported infectious events by PT (RUX vs BAT arm) were pneumonia (10.9% and 12.7%), upper respiratory tract infection (8.5% and 8.2%), urinary tract infection (6.7% and 3.2%), nasopharyngitis (6.1% and 3.8%), BK virus infection (5.5% and 1.3%), CMV infection (5.5% and 8.2%), influenza (4.8% and 3.8%), and conjunctivitis (4.8% and 2.5%).

- AE pneumonia up to up to Cycle 7, Day 1 was seen in 11% vs 12% in the respective treatment arm (RUX vs BAT) and CMV infections in 9.1% in the RUX arm vs. 10.8% in the BAT arm. Other PT CMV infections were also separately reported.

3.5. Uncertainties and limitations about unfavourable effects

Assessment of the safety profile of ruxolitinib and the crude incidence of AEs is challenging in a setting of immune compromised patients with severe underlying disease and concomitant use of corticosteroids and CNI bringing uncertainties with respect to the assessment of treatment-related AEs.

The latest updated OS data are 25 Jun 2021 for REACH-3 (22.4% vs 22.0% for the RUX and the BAT arm respectively). The allowance of cross-over in both studies creates a bias towards unity with regards to the impact of treatment on OS.

A proportion of treatment related AEs in REACH2 were not resolved at data cut-off, however, this was based on AEs that were not recovered/not resolved in patients who discontinued and who have finished their 30-day safety follow-up. In both REACH 2 and REACH 3 the most frequent AEs that were not resolved were cytopenias, all other SAEs were reported infrequent and with no specific pattern and not leading to new safety concerns.

The requested indication includes adolescents. The AE profile was largely similar to adults with some exceptions. Nevertheless, RUX treatment is studied in a limited dataset (5 patients with aGVHD, 4 patients with cGvHD).

A significant proportion of subjects in study REACH 2 (19.1%) and REACH 3 (23%) discontinued RUX treatment due to physician decision (thus not classified as discontinuation due to lack of efficacy or due to an AE), it is unclear what the primary reason for discontinuation is and further information was not captured in the CRFs.

3.6. Effects Table

Effects Table for Jakavi (ruxolitinib)

Effect	Short description	Unit	Treatment Ruxolitinib 10 mg BID	Control BAT	Uncertainties / Strength of evidence	References
Favourable Effects REACH-2 aGvHD: Data cut offs: 25-Jul-2019 (Primary analysis, PA), 06-Jan-2020 (second analysis, SA)						
			N=154	**N=155**	One-sided p-values	

Effect	Short description	Unit	Treatment Ruxolitinib 10 mg BID	Control BAT	Uncertainties / Strength of evidence	References
ORR by INV, at **Day 28**	Overall response rate (CR + PR)	%	62.3 95% CI (54.2, 70.0)	39.4 95% CI (31.6, 47.5)	OR: 2.64: 95% CI (1.65, 4.22). p<0.0001	Primary endpoint (PA)
Durable ORR at **Day 56**	responders only	%	39.6 95% CI (31.8, 47.8)	21.9 95% CI (15.7, 29.3)	OR: 2.38: 95% CI (1.43, 3.94) p=0.0005.	Key secondary endpoint (PA)
BOR Best overall response by Day 28	ORR at any time point up to and incl. Day 28	%	81.8 95% CI (74.8, 87.6)	60.5 95% CI (52.5, 68.4)	OR*: 3.07: 95% CI (1.80, 5.25) p<0.0001	Other secondary endpoints (PA) goes for all endpoints below
DoR Duration of response (Median)	Composite endpoint [a]	Days	163 range: 22.0 to 623.0	101 range: 10.0 to 456.0	NS*	Only pat who had a CR or PR at or before Day 28 (SA, goes for all endpoints below)
OS (Median)	Overall Survival	months	10.71	5.82	HR: 0.83 (0.62, 1.13) P=0.2648 NS*	
EFS Event free survival (Median)	Composite[b] endpoint	months	8.18	4,17	HR: 0.80 (0.60, 1.08) NS*	
FFS Failure free survival (Median)	Composite[c] endpoint	months	4.86	1.02	HR: 0.49, 95% CI: 0.37, 0.63) Descriptive	
Favourable Effects REACH-3 cGvHD: **Data cut-off: 08-May-2020 Primary Analysis (PA)**						
			N=165	**N=164**		
ORR at Cycle 7 Day 1	(CR + PR)	%	49.7 95%CI (41.8, 57.6)	25.6 95%CI (19.1, 33.0)	OR: 2.99 95% CI (1.86, 4.80) p<0.0001	Primary endpoint (PA)
FFS Failure free survival Median	Composite endpoint	months	NE (18.6, NE)	5.7 (5.6, 6.5)	HR: 0.370 95% CI (0.268, 0.510) p<0.0001.	First Key secondary endpoint (PA)
TSS PRO based on the total symptom score,	Based on modified Lee symptom scale		24.2 (17.9, 31.5)	11.0 (6.6, 16.8)	OR: 2.62 (95% CI: 1.42, 4.82) p = 0.0011	Second Key secondary endpoint (PA)
BOR up to Cycle 7 Day 1		%	76.4 (69.1, 82.6)	60.4 (52.4, 67.9)	2.17 (95% CI: 1.34, 3.52)*	
OS		n/N	37/165	36/164	HR: 0,956 ((95% CI: 0.604, 1.512)*	(updated 25 Jun 2021
DOR		months	NE (20.2,	6.2 (4.7,	Descriptive	Measured only

Effect	Short description	Unit	Treatment Ruxolitinib 10 mg BID	Control BAT	Uncertainties / Strength of evidence	References
(Median)			NE)	13.3)		pat who had a CR or PR at or before Cycle 7, Day (PA)
Unfavourable Effects REACH-2 aGvHD						
	Up to SA data cut-off	%	N=154	N=155		
Any AE			99.3	98.7		
AE Gr ≥ 3			91.4	87.3		
AE leading to dose adjustment /interrupt.			54.6	13.3		
AE leading to discont.			27	9.3		
TEAEs ≥ 40 %	Infections Blood and lymphatic disorder G-I disorder Investigations General disorder Metabolism /nutrition	%	81.6 71.1 63.8 58.6 53.9 50.0	71.3 51.3 50.0 46.0 50.0 53.3		
AEs by PT >20%	 Anemia Thrombocytopenia Neutropen. CMV infect. Oedema peripheral Hypokalemia Pyrexia Platelet count decreased	Any/Gr ≥3 (%)	 40.1/35.5 36.8/33.6 24.3/21.7 30.9/9.2 24.3/2.0 22.4/9.9 22.9/2.0 20.4/17.8	 32.0/24.0 20.7/16.7 14.7/12.0 26.7/12.0 21.3/2,0 21.3/2.0 21.3/2.0 16.0/15.3		
Treatment related AEs (suspected)	Blood and lymphatic disorder Investigations	All/Gr≥ 3 (%)	40.1/36.2 28.9/22.4	10.0/8.7 12.7/9.3		

Effect	Short description	Unit	Treatment Ruxolitinib 10 mg BID	Control BAT	Uncertainties / Strength of evidence	References
	Infections GI disorder Nerv. System disorders		26.3/19/1 13.8/5.3 7.9/2.6	13.3/10.7 3.3/1.3 4.0/1.3		
AESI Gr≥ 3 >10%	Trombocytopenia Leucopenia Anemia Infection (excl Tb)** CMV infection (any grade/ Gr≥3) Sepsis/chock Pneumonia Opportun. Infections Other infections Bleeding events Hypertension		50.7 42.8 36.2 52.0 30.9 (9.2) 20.4 11.8 7.9 27.0 12.5 13.8	32.0 27.3 25.3 47.3 26.7 (12.0) 18.0 14.0 4.7 24.0 7.3 12.7		
SAEs		%	66.4	53.3		
SAE Gr≥ 3	 Sepsis Septic chock	%	63.8 5.3 4.6	50.0 2.7 2.7		
Fatal SAEs		%	21.7	21.3		
Deaths up to DCO			53.9	57.3		
Deaths (on treatment)		%	28.3	24.0		
Deaths due to aGvHD		%	13.8	14.0		
Deaths (suspect treatment related)		N	10	4		
Unfavourable Effects REACH-3 cGvHD						
	During the Main treatment period	%	N=165	N=164		
Any AE			98.2	92.4		
AE Gr ≥ 3			66.1	58.9		
AE leading to dose			43.0	18.4		

Effect	Short description	Unit	Treatment Ruxolitinib 10 mg BID	Control BAT	Uncertainties / Strength of evidence	References
adjustment /interrupt.						
AE leading to discont.			20.6	8.9		
TEAEs ≥ 40 % By SOC	Infections Investigations Blood and lymphatic disorder Metabolism /nutrition Respiratory etc GI disorder General disorders	Any/Gr ≥3	71.5/26.1 56.4/29.7 46.1/28.5 44.2/21.2 43.6/12.1 42.4/6.1 42.4/7.3	65.8/22.2 34.8/17.7 24.1/16.5 38.0/15.2 34.8/9.5 42.4/9.5 36.7/6.3		
AEs by PT >10%	Anemia Pyrexia ALT increase Hypertension Blood creatin. Increase Diarrhoea Pneumonia Cough Neutropenia Fatigue Trombocytopenia Nausea Dyspnoea Platelets decrease AST increased GammaGT increase	Any/Gr ≥3	32.1/15.2 20.0/3.0 17.6/6.1 17.6/5.5 15.8/0 15.8/1.2 15.8/13.3 13.9/0.6 13.9/12.1 12.1/0.6 12.1/11.5 11.5/0 10.9/1.8 10.9/6.1 10.3/1.8 10.3/7.3	13.9/7.6 10.8/1.3 4.4/0 13.3/7.0 4.4/0.6 15.8/1.9 13.3/10.1 8.9/0 5.1/3.1 10.1/1.9 8.9/5.9 7.0/1.9 7.0/1.9 7.0/5.7 2.5/0.6 3.8/2.5		
Treatment related AEs (susp) >5%	Anemia Neutropen. ALT increased Trombocytopenia Platelet count decreased Pneumonia AST increased Blood creatin.	Any/Gr ≥3	23.6/10.3 10.9/8.5 10.3/4.2 9.1/8.5 6.1/3.0 7.3/6.7 5.5/0.6 5.5/0	3.2/1.3 3.271.3 0.6/0 3.8/1.3 1.9/1.3 3.2/2.5 0.6/0 0.6/0		

Effect	Short description	Unit	Treatment Ruxolitinib 10 mg BID	Control BAT	Uncertainties / Strength of evidence	References
	increased					
AESI >10%	Infection (excl Tb)**	Any/Gr ≥3	72.1/26.7	65.8/22.2		
	Pneumonia		23.6/18.2	17.7/14.6		
	Opportun. Infections		12.7/3.6	13.3/3.8		
	UTI		10.9/1.8	9.5/1.9		
	CMV infection (any grade)		5.5/1.2	8.9/0		
	Other infections		58.8/10.9	53.2/10.8		
	Anemia		32.1/15.2	13.9/7.6		
	Leucopenia		23.6/18.8	14.6/11.4		
	Trombocytopenia		23.0/17.6	15.8/11.4		
	Bleeding events		12.7/3.0	16.5/3.2		
	Elevated ALT/AST		20.6/8.5	7.6/0.6		
	Lipid abnormal		20.6/7.3	14.6/7.0		
	Hypertension		18.2/6.7	13.3/7.0		
SAEs		Any/Gr ≥3 (%)	43.6/40.0	39.9/36.1		
>5%	Pneumonia		12.7/12.1	8.9/7.6		
	Pyrexia		6.7/3.0	2.5/1.3		
Fatal SAEs			9.1	6.3		
Deaths			18.8	16.5		
Deaths (on treatment)			9.7	5.7		
Deaths due to cGvHD			8.5	4.4		

*Not corrected for multiplicity

** No Tb cases reported

Abbreviations: OR=Odds ratio, NS=not significant, HR= Hazard ratio, NE= not estimated, NRM=non-relapse mortality, PA=primary analysis, SA=second analysis, IA=interim analysis

Notes:
[a] **i**) Progression of aGvHD **ii**) addition of systemic therapy for aGvHD after D28.
Competing risk: death without prior observation of aGvHD and onset of cGvHD
[b] **i**) Hematol disease relapse **ii**) graft failure or death due to any cause
[c] **i**) hematologic disease relapse/progression **ii**) NRM **iii**) addition of new systemic aGvHD treatment.
Competing risk: cGvHD
[d] **i**) relapse or recurrence of underlying disease or death due to underlying disease **ii**) non-relapse mortality **iii**) addition or initiation of another systemic therapy for cGvHD.

3.7. Benefit-risk assessment and discussion

3.7.1. Importance of favourable and unfavourable effects

The applicant has selected objective response rate as primary endpoint in each of its studies on acute and chronic GvHD. While this is anticipated to confer symptomatic benefit, the applicant claims it is a surrogate for long term outcomes. The applicant has provided statistically robust evidence of the effects of ruxolitinib on this endpoint as defined. However, intensification of immunosuppressive treatment was considered failure to reach an objective response in the primary analysis. Thus, it remains unclear to what extent this is due to an overall increase in regimen efficacy, and to what extent this is due to sparing of other drugs through substitution. In either case, this is considered clinical benefit.

The safety profile of Jakavi is similar to what has been seen in the previously approved indications and is dominated by cytopenias and increased risk of infection. There was no apparent increase in the risk of relapse or progression of malignancy.

No effects on OS have been shown, which may be impacted by the cross-over option which creates bias towards unity.

The proposed indication does not specify GvHD type or grade and allows treatment of patients that are first line steroid refractory as well as patients that have received multiple systemic therapies for GvHD. A reference to section 5.1 in the indication has therefore, been provided with regard to detailed information of the studied population.

Overall, despite one critical finding and several major findings identified during the GCP inspection which showed that on several key aspects the processes were not robust for a sound conduct of the study and/or failing to prevent deviations from ICH-GCP, it was the recommendation of the inspectors that the data of the REACH-2 and REACH-3 clinical trials were of sufficient quality to be used for evaluation and assessment of the application, though the inspectors deferred to the assessors for the final conclusion on the impact of the lacking documentation on best available therapy (BAT) accountability in REACH- 2 and REACH-3.

Notably, the study protocols for the REACH-2 and REACH-3 clinical trials explicitly warranted drug accountability for ruxolitinib only, but according to the most conservative assessment of the inspection team, for 12% out of 150 trial participants for REACH-2 and 41% out of 159 patients for REACH-3, BAT treatment compliance was insufficiently monitored. For a*cute GvHD* the approximately 12% of patients without drug accountability for BAT in REACH-2 still allows sufficient interpretation of clinical study results, allowing for a positive B/R.

Additional exploratory analyses to assess the primary efficacy response of the different types of BAT used in the *chronic GvHD* study REACH-3 showed for each independent BAT a lower ORR than the response rate reported in patients treated with RUX, regardless of the administration route and/or clinical setting, i.e. in hospital / at home. Furthermore, the in-hospital BATs, which have high drug accountability (due to the setting), and the at-home BATs showed similar ORR per BAT. These exploratory results together are suggestive of a limited (if any) impact of the potential lower BAT compliance (worst case) in the at-home setting.

As OS is multifactorial, for which BAT compliance is only one parameter of interest, it remains difficult to establish whether potential non-compliance in the BAT arm has substantially influenced OS outcomes. It is reassuring that with the latest OS update no sign of detriment was observed.

Finally, even if outcomes would have been comparable to BAT, ruxolitinib could still be acceptable as new treatment modality for chronic GvHD. Therefore, despite some remaining uncertainties on the impact of potential non-compliance on the efficacy results due to the exploratory nature of the additional analyses, the B/R of ruxolitinib is considered positive in chronic GvHD as well.

3.8. Conclusions

The B/R of Jakavi is positive in acute GvHD and chronic GvHD. Final OS data from REACH-3, including modelling of the impact of cross-over, should be provided as a post marketing recommendation.

4. Recommendations

Outcome

Based on the review of the submitted data, the CHMP considers the following variation acceptable and therefore recommends the variation to the terms of the Marketing Authorisation, concerning the following change:

Variation accepted		**Type**	**Annexes affected**
C.I.6.a	C.I.6.a - Change(s) to therapeutic indication(s) - Addition of a new therapeutic indication or modification of an approved one	Type II	I and IIIB

Extension of indication to include treatment of patients with acute and chronic GvHD aged 12 years and older who have inadequate response to corticosteroids or other systemic therapies for Jakavi; as a consequence, sections 4.1, 4.2, 4.4, 4.8. 5.1 and 5.2 of the SmPC are updated. The Package Leaflet is updated in accordance. Version 14.0 of the RMP has also been submitted. In addition, the Marketing authorisation holder (MAH) took the opportunity to update the list of local representative for The Netherlands in the Package Leaflet.

The variation leads to amendments to the Summary of Product Characteristics and Package Leaflet and to the Risk Management Plan (RMP).

Amendments to the marketing authorisation

In view of the data submitted with the variation, amendments to Annex(es) I and IIIB and to the Risk Management Plan are recommended.

Paediatric data

Furthermore, the CHMP reviewed the available paediatric data of studies subject to the agreed Paediatric Investigation Plan EMEA-000901-PIP03-16-M01 and EMEA-000901-PIP04-17-M01and the results of these studies are reflected in the Summary of Product Characteristics (SmPC) and, as appropriate, the Package Leaflet.

5. EPAR changes

The EPAR will be updated following Commission Decision for this variation. In particular the EPAR module "*steps after the authorisation*" will be updated as follows:

Scope

Please refer to the Recommendations section above.

Summary

Please refer to Scientific Discussion 'Jakavi-H-C-2464-II-0053'