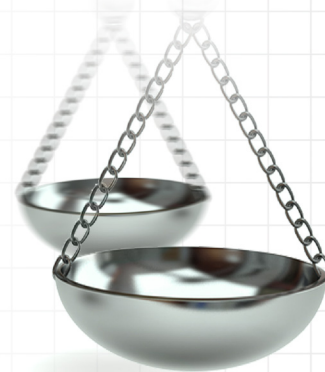




PHEDRA: using real-world data to analyze treatment patterns and ibrutinib effectiveness in hematological malignancies

Jamie Garside^{*1}, Nollaig Healy², Hervé Besson³, Ruben Hermans⁴, Finlay MacDougall⁴, Damien Lestelle⁴, Joris Diels³ & Wafae Iraqi⁵

¹Janssen EMEA HEMAR, High Wycombe, UK

²Janssen-Cilag EMEA Medical Affairs, Dublin, Ireland

³Janssen EU HEMAR Statistics & Modelling, Beerse, Belgium

⁴QuintilesIMS RWE Solutions, London, UK

⁵Janssen Pharmaceuticals, Paris, France

* Author for correspondence: Tel.: +44 (0) 149 465 8410; Fax: +44 (0) 782 528 0450; jgarside@ITS.JNJ.com

Aim: PHEDRA (Platform for Haematology in EMEA: Data for Real World Analysis) is a unique, noninterventional project based on secondary data collection from real-world (RW) patient-level (health record) databases to understand treatment patterns in hematological malignancies. It compares ibrutinib's effectiveness with alternative treatments using RW data (RWD) and randomized clinical trials data. **Materials & methods:** RWD are cleaned, validated, harmonized into a Common Data Model, and analyzed statistically alongside randomized clinical trial data. Treatment outcomes include overall and progression-free survival. **Results:** To date, RWD (four databases) are available for 2840 patients in three indications, collected between 1990 and 2017. **Conclusion:** PHEDRA is an innovative approach to generate evidence to inform optimal treatment decisions in RW settings.

First draft submitted: 20 June 2017; Accepted for publication: 14 July 2017; Published online: 2 August 2017

Keywords: • comparative effectiveness research • hematology • methodology

Randomized clinical trials (RCTs) enable medicines to be rigorously evaluated in a controlled environment, through the application of inclusion/exclusion criteria for patient entry into the trial and randomization of treatments administered [1]. However, they do not establish the comparative effectiveness of drugs in real-world (RW) populations, which are often more heterogeneous than trial populations and where treatment is often subject to less close monitoring and frequency of data collection [1–3]. There remains insufficient knowledge and limited published literature on routine clinical practice, treatment pathways and outcomes in the management of patients with hematological malignancies. This is needed to guide optimal therapeutic decisions in patients with cancer at different stages of disease, particularly as the number of available treatments and possible drug combinations multiply [4].

There is increasing recognition of the value of using real-world data (RWD) alongside RCT data (i.e., historical controls) to generate comparative effectiveness research on new therapies [2,5]. Indeed, with advances in cancer therapy and a multiplicity of comparators/standards of care depending on local routine practice (making it impractical or impossible to obtain data on all potentially relevant comparisons in an RCT setting), the need for RWD to educate regulators, payers, healthcare providers, healthcare decision makers and patients about treatment patterns and outcomes has never been greater [2]. RWD can further demonstrate the comparative effectiveness of a treatment once available on the market as well as inform clinical drug development, for example, by providing information on existing therapies and the profile of patients with poor outcomes who may have specific unmet needs. Although RWD cannot replace RCT data with regard to efficacy and safety, it can challenge and improve understanding of the 'standard of care' [1]. RWD can provide useful outcomes information across multiple relevant

regimens, which can be compared with data from single-arm trials or complement the findings of head-to-head trials that were originally designed with only one comparator in mind, while other comparators may be relevant, especially based on the different clinical practices in respective countries. RWD, therefore, offers the advantages of demonstrating comparative treatment effectiveness in the clinical practice setting and confirming RCT results in real-life patient populations [2].

Ibrutinib is a treatment for B-cell malignancies, approved in the EU [6], the USA [7] and elsewhere to treat chronic lymphocytic leukemia (CLL), Waldenström's macroglobulinemia (WM) and relapsed/refractory mantle cell lymphoma (MCL). It is also indicated for marginal zone lymphoma in the US [7]. Despite the availability of a range of RCTs investigating the treatment of these diseases, there is still a lack of comparative evidence on treatment selection, duration of treatment and disease outcomes across different lines of therapy. Since ibrutinib is relatively new to the market (first US FDA approval in 2013 and European Medicines Agency [EMA] approval in 2014), head-to-head comparisons of ibrutinib versus other widely used treatments in CLL, MCL and WM are limited (outside the RCTs used for approval) [8–10], and although there are growing RWD for ibrutinib, comparative effectiveness studies using RWD only are yet to reach optimal data volume and quality.

Therefore, Janssen Pharmaceutica NV developed an RWD project known as the Platform for Haematology in EMEA: Data for Real World Analysis (PHEDRA). PHEDRA has two main objectives: to understand current RW treatment practice and outcomes in CLL, MCL and WM, and to compare the effectiveness of ibrutinib therapy (based on RCT data and RWD) with a range of other treatment regimens routinely used in clinical practice and documented in RW databases. The unique approach for PHEDRA is the subject of this manuscript; the aims are to describe the main steps taken to validate and analyze the independent, existing RWD, and to provide the necessary background on the methodology of how the comparative evidence is generated and how the PHEDRA results can be interpreted (taking into account the strengths and inherent limitations related to RWD [11]).

Materials & methods

PHEDRA is a noninterventional, secondary-use project which collaborates with the owners of existing databases of electronic health records in the Europe, Middle East and Africa (EMEA) region and gathers deidentified patient-level data on CLL, MCL and WM centrally. The project is planned to last for 8 years with regular analyses on longitudinal retrospective patient data from the selected databases. Identified, independent sources of RWD (existing registries/databases) are qualitatively and quantitatively assessed, to ensure inclusion of RWD that capture the necessary clinical variables (relevant outcomes) to answer the two main objectives.

The PHEDRA methodology is divided into two main steps. First, external data are extracted from the different participating clinical centers (databases), cleaned and then harmonized into one Common Data Model (CDM). The second step involves the data analysis. An overview of these steps is provided in [Figure 1](#).

Data extraction

The RWD sources are pre-existing, and all data are retrospective when extracted. PHEDRA aims to extract data from the different RWD sources annually; each RWD extract comprises patients included in the previous dataset (longitudinal follow-up) as well as new patients treated since the last extract. Although the data extraction requirements do not interfere with the processes of data capture in routine clinical practice, efforts are made to improve the data quality and completeness of the individual databases.

Data deidentification

Each RWD extract includes only deidentified patient-level data. This process is a legal requirement to ensure patient privacy and is performed by the database owners at their respective sites prior to data transfer. Upon receiving these data, it is confirmed that the deidentification has been performed. This occurs before aligning the data format and starting the data harmonization process.

Data harmonization into a CDM

Data integration of the RW databases participating in PHEDRA is driven by a CDM, which has a relational table structure, a variable format and set of reference libraries, and enables the source data in its native form to be conformed to a common standard. The subsequent data handling processes can be defined through the following components: cleaning and harmonizing data, data validation and enhancement, and dealing with incomplete data.

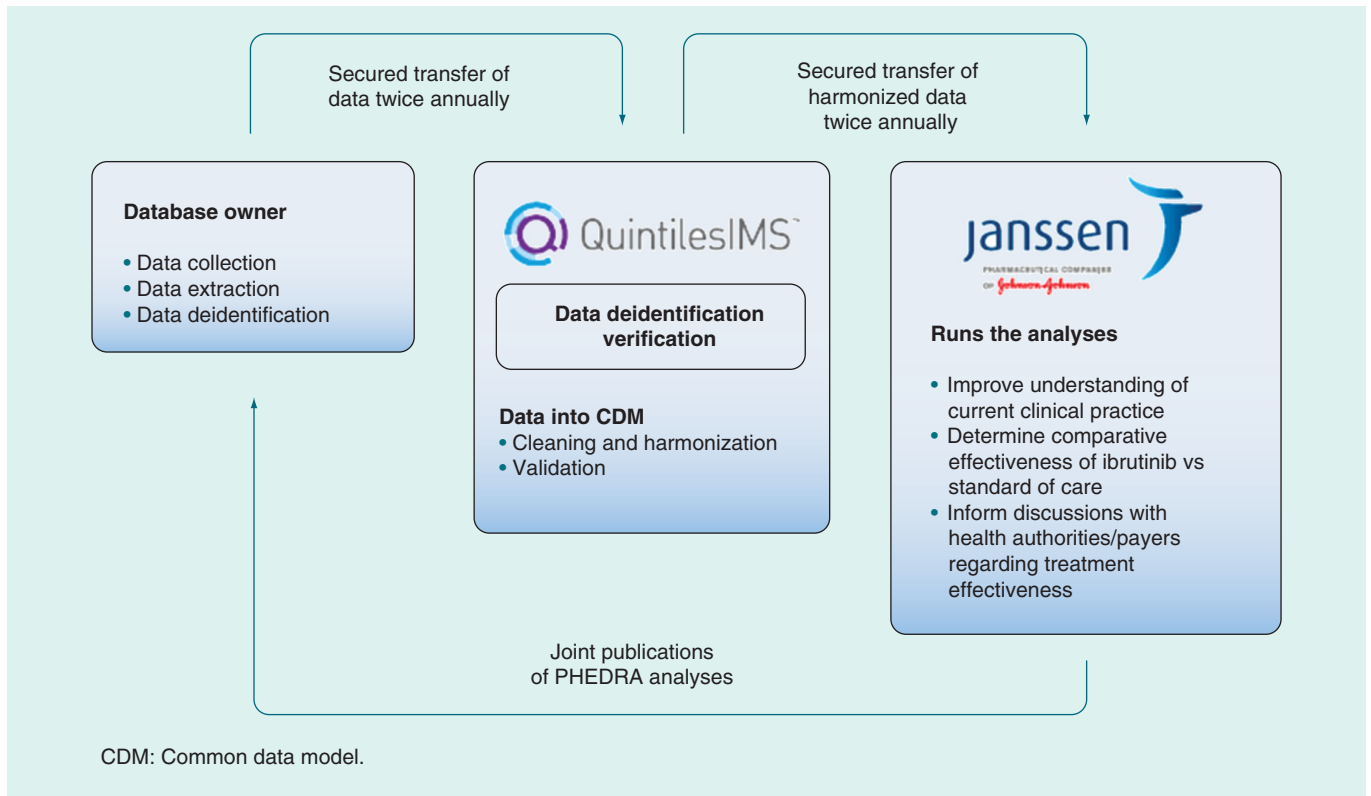


Figure 1. PHEDRA overview and methodology.
CDM: Common Data Model.

Cleaning & harmonizing of data

With the aim of improving the consistency of the different source datasets, the data are first cleaned and then harmonized into a normalized format. The process of harmonization encompasses the technical process of mapping distinct data elements from the source data to the CDM [12]. Depending on the structure of the source dataset (i.e., the format in which the variables have been entered into the database), the number of data cleaning and harmonization steps can vary. For example, regarding a patient's Rai disease stage, some databases will have their entries registered as 0 (lymphocytosis only) to IV (lymphocytosis with thrombocytopenia), whereas others will have it recorded in numerical values as 1 (lymphocytosis only) to 5 (lymphocytosis with thrombocytopenia). The CDM will create a common format, for example, all Rai staging will be entered as 0–IV.

Data validation & enhancement

Data validation and enhancement ensure the completeness, accuracy, internal logical consistency and suitability of datasets for subsequent statistical analysis, and is conducted before, during and after the harmonization process.

During data harmonization, validation checks are carried out to assess quality of the source data. After the data have been harmonized to the CDM format, any inconsistencies, missing data fields and values outside of expected ranges are logged and shared with the RW database owners, to address issues and improve data quality and completeness.

Dealing with incomplete data

All missing data fields (including baseline characteristics) are queried with the database owners, to minimize the level of incomplete data. Data fields that cannot be completed by database owners through reference to existing secondary data are categorized as 'missing'.

In addition to the CDM, two main types of adjustments need to be applied to make the data ready to be analyzed in comparison with RCTs. Firstly, selection of variables available in both the RCTs and the RW cohorts is harmonized by applying the definition from the RCT. For instance, the defined cut-off for the 'Bulky Disease'

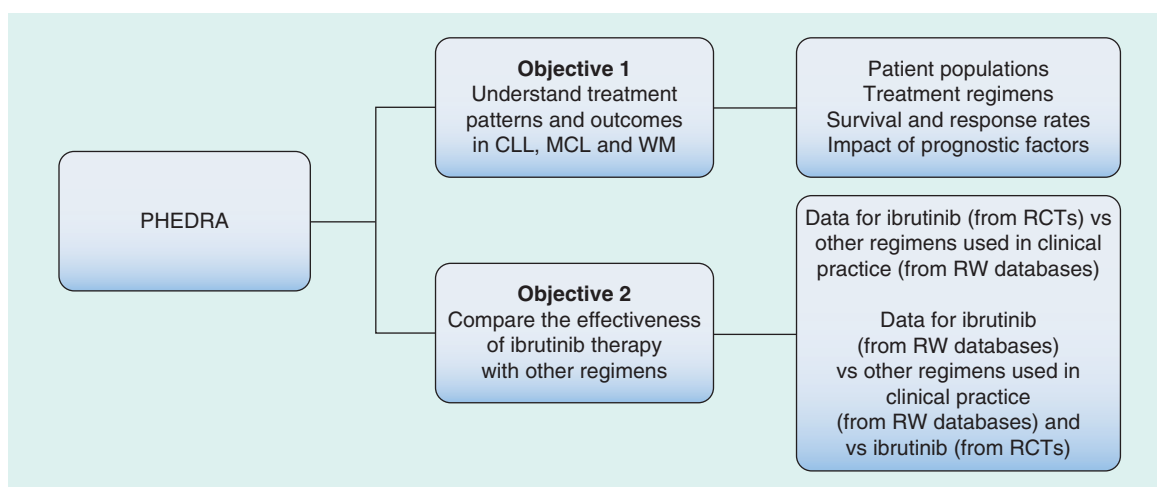


Figure 2. PHEDRA objectives.

CLL: Chronic lymphocytic leukemia; MCL: Mantle cell lymphoma; PHEDRA: Platform for Haematology in EMEA: Data for Real World Analysis; RCT: Randomized controlled trial; RW: Real world; WM: Waldenström's macroglobulinemia.

variable applied to RW cohorts is the same cut-off as defined in the RCT protocol. Secondly, those patients from the RW cohorts who appear to be enrolled in RCTs and/or treated with experimental regimens are excluded.

Analyses

The harmonized data will be analyzed by using statistical and descriptive techniques to meet the PHEDRA objectives, as summarized in Figure 2.

RW treatment & outcomes (objective 1)

To meet the first objective, specific analyses are conducted on the data, to identify and evaluate the following for the CLL, MCL or WM cohorts:

- The patient population including detailed patient and disease characteristics, summarized using descriptive statistics;
- The treatment pathway of patients from diagnosis until death or end of care;
- Progression-free survival (PFS) or time to progression (TTP) and time to next therapy (TTNT) by patient and disease characteristics and by line of treatment; PFS is defined as the time between the date of treatment initiation and the date of disease progression or date of death due to any cause, whichever occurs first;
- Overall survival (OS) from diagnosis and from each line of treatment; OS is defined as the time from date of diagnosis or date of treatment initiation to date of death from any cause;
- Response rates, time to response and response duration for individual treatment regimens;
- The relationship between type and line of therapy, response, TTP/PFS and TTNT and OS for different treatment regimens;
- The impact of prognostic factors, including cytogenetic abnormalities, on response, TTP/PFS and OS.

Subgroups and treatment exposure are predefined in the disease-specific statistical analysis plans prior to the first analysis. Descriptive statistics (n, mean, SD, median and range) are provided for treatment duration.

Comparative effectiveness analyses (objective 2)

Hypothesis & analysis populations

The comparative effectiveness analyses are conducted for ibrutinib compared with other RW treatment regimens. The hypothesis is that ibrutinib significantly prolongs survival outcomes in patients within the targeted indications, compared with any RW standard-of-care (physician's choice) regimen included in PHEDRA. This hypothesis is tested using an adjusted comparison of ibrutinib patient-level data from RCTs and RW cohorts with other RW regimens within PHEDRA, and the analyses are conducted on the intention to treat populations.

Methodological considerations/challenges

With the project being retrospective, no sample size calculation is performed as the size of the population is determined by the size of the available PHEDRA databases. For comparisons with RCT data, only patients in the RW databases who meet the RCT inclusion/exclusion criteria are included. However, all other patients can be included in the analyses when simply describing the outcomes of a cohort, or when making comparisons within a cohort. PHEDRA databases are studied either separately or by pooling data at the individual patient-level across the respective databases, providing comparability of outcome variables and baseline characteristics. Pooling databases increases the number of patients analyzed, enhances the ability to interpret subgroup analyses (e.g., to study specific regimens) and captures the consistency of results across different clinical settings. Analyses using more than one RW database take into account whether the estimates are considered homogeneous or heterogeneous across databases by using either a fixed effect model or random effect model, respectively.

The unit of observation in these analyses is the treatment line, rather than the patient. When patients move into further treatment lines after their first-line therapy, these further treatment lines are included as separate observations, and so a patient can contribute to multiple observation units if they receive multiple treatment lines. For each line of therapy, patient characteristics are collected at the initiation of each treatment line, reflecting the corresponding baseline status of the patient. The multiple use of the entire treatment sequence of one patient is based on the following rationale: consider a relapsed/refractory patient treated in second line with regimen A, who progressed and was subsequently initiated on another regimen in third line. The outcome of this same patient in the third-line setting can be representative of any other third-line patients with a similar treatment history and so these data are used within the analysis; disregarding the data on this same patient in third or later lines would be a missed opportunity to use the information provided by the databases. However, it needs to be acknowledged that these multiple observations are of course correlated because they refer to the same patient, which goes against traditional assumptions regarding the independence of observations in statistical analyses. This is accounted and controlled for by using the robust sandwich estimate for the covariance matrix [13,14], which makes standard errors and confidence intervals (CIs) around point estimates somewhat broader, compared with the situation where all of these observations would have come from different patients and would have been truly independent. A similar principle of including the same patient multiple times, each of them with a different point of follow-up, was proposed by Hernan *et al.*, who considered this approach more efficient from a statistical standpoint, provided that ‘appropriate adjustment of the usual variance estimator’ is implemented [15].

As the current focus of the PHEDRA project is to compare ibrutinib-treated patients from RCTs to RW treatments, only the treatment line involving ibrutinib from RCTs [8,10,16] is included in the analyses performed.

Outcome analyses

To account for noncomparability of patient populations due to lack of randomization in RW databases, a multivariate Cox proportional hazard regression model is used to derive adjusted hazard ratios (HRs) for the relative treatment effect of ibrutinib versus the physician’s choice in the RW cohort, including prognostic factors for PFS and OS as covariates. The list of baseline characteristics included in the model covers a range of clinically relevant predictors for PFS/OS, which are available for both the RW cohort and the RCT, such as age, sex, disease stage (based on Binet/Rai), the presence of deletion 17p or 11q and line of therapy. The adjusted HRs for treatment effects and prognostic covariates are graphically presented using forest-plots, displaying the point estimate and the 95% CI. For the analysis of PFS, patients who are still progression free and alive at the time of the data cut (or lost to follow-up) are censored at the last day the patient was known to be progression free. Similarly, for OS, patients who are still alive at the time of the data cut (or lost to follow-up) are censored at the last day the patient was known to be alive. OS is analyzed using the same analysis methods as used for PFS.

Descriptive and comparative analyses are performed using appropriate statistical methodology, including Kaplan–Meier analyses. Data are stratified by treatment and/or prognostic baseline characteristics, and Kaplan–Meier plots for these stratified analyses are graphically represented. The data from the RW databases and the related RCTs are combined. The following steps are followed to generate the predicted survival curves for both the end points of PFS and OS. Based on the multivariate model, predicted survival curves are estimated by patient, using patient-specific baseline characteristics as covariates. A first survival curve represents the predicted patient-specific survival under treatment as observed (representing the outcome as treated). A second survival curve represents the predicted survival under the scenario that patients are treated with ibrutinib rather than with the treatment they actually

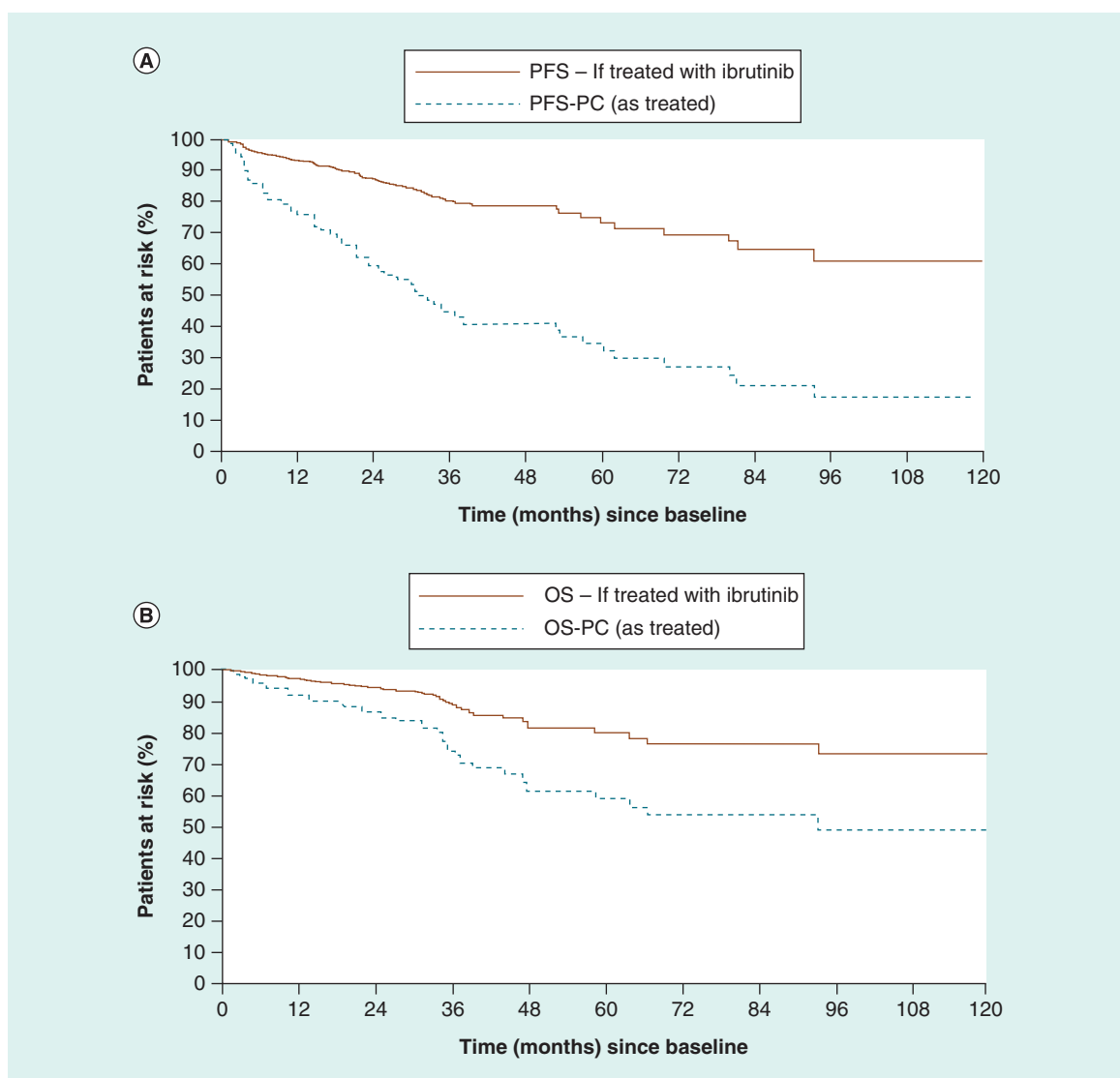


Figure 3. PHEDRA predicted progression-free survival (A) and overall survival (B) Kaplan-Meier curves for treatment-naïve CLL patients from Lyon-Sud if treated with ibrutinib versus physician's choice in real-world. CLL: Chronic lymphocytic leukemia; RW: Real world. Data taken from [18].

received; these predictions are based on the same statistical model as used for the first predicted survival curves, but incorporate the improved outcome related to ibrutinib, by adding the parameter estimate for ibrutinib treatment versus physician's choice to the Cox proportional regression model. As a last step, the mean survival for the entire cohort over time is generated for both scenarios (as treated in RW vs if the cohort had received ibrutinib treatment) by simply averaging predicted survival rates over time across patients. This results in two predicted survival curves (as treated in RW vs if the cohort had received ibrutinib treatment) for the population. The difference between both curves represents the adjusted HR for ibrutinib relative to physician's choice of treatment. The two survival curves are illustrated on the same plot in the publications (Figure 3).

Future analyses will also explore the extent to which the harmonized data will allow evaluation of ibrutinib versus other regimens according to observed response rates (overall response, complete response, partial response, etc.). This will be analyzed using multivariate logistic regression models when data are available in the databases.

Results & discussion

PHEDRA is designed to enable comparisons of patient populations, treatment configurations and outcomes over multiple lines of therapy and within one or across several RW databases. The results from these analyses allow, for the first time, both clinicians and payers to interpret the comparative effectiveness of ibrutinib versus RW treatments commonly used in routine clinical practice.

To date, the four RW databases selected for the PHEDRA project have collected information on a combined population of patients with CLL of $n = 2504$ from four countries between 1990 and 2017. The initial data extractions for MCL and WM were not available at the time of manuscript development.

Analyses on two RW databases in patients with CLL were presented at major international hematology congresses in 2016 and 2017 [17–19]. The first presentation evaluated the comparative effectiveness of ibrutinib versus the RW treatments by conducting an adjusted comparison of patient-level data from one clinical trial (RESONATE™) with a single RW cohort from the CLLEAR database in the Czech Republic [17]. In relapsed and refractory CLL patients, the analysis showed that for PFS, the adjusted HR of ibrutinib versus all treatments in CLLEAR was 0.10 (95% CI: 0.06–0.16; $p < 0.001$). For OS, the adjusted HR of ibrutinib versus all treatments in CLLEAR was 0.14 (95% CI: 0.08–0.26; $p < 0.001$).

The second presentation evaluated the comparative effectiveness of ibrutinib versus the RW treatments by conducting an adjusted comparison of patient-level data from two clinical trials (RESONATE™ and RESONATE-2™) with a treatment-naïve and relapsed/refractory RW cohort from the Lyon–Sud database in France [18]. In treatment-naïve CLL patients, the analysis showed that for PFS, the adjusted HR of ibrutinib versus all treatments in Lyon–Sud was 0.26 (95% CI: 0.14–0.49; $p < 0.001$). For OS, the adjusted HR of ibrutinib versus all treatments in Lyon–Sud was 0.40 (95% CI: 0.17–0.97; $p = 0.042$). The predicted survival curves are shown in [Figure 3](#).

In relapsed/refractory CLL patients, the analysis showed that for PFS, the adjusted HR of ibrutinib versus all treatments in Lyon–Sud was 0.18 (95% CI: 0.13–0.26, $p < 0.001$). For OS, the adjusted HR of ibrutinib versus all treatments in Lyon–Sud was 0.27 (95% CI: 0.16–0.44, $p < 0.001$).

The third presentation details the results of an analysis pooling more than one RW database for patients with treatment-naïve CLL [19]. Adjusted comparison of data from RESONATE-2™ with RW data from CLLEAR and Lyon–Sud demonstrated that for PFS, the HR of ibrutinib versus all treatments was 0.27 (95% CI: 0.16–0.45; $p < 0.001$), and for OS, the HR was 0.40 (95% CI: 0.19–0.82; $p = 0.0120$) for treatment-naïve CLL patients ≥ 65 years without del17p. The first analysis in MCL and in WM will be performed and submitted for publication later in due course.

The results published to date demonstrate the relative comparative effectiveness of ibrutinib versus RW treatments, in both treatment-naïve and relapsed and refractory CLL patient populations. The adjusted HRs for PFS and OS across the separate databases and analyses all suggest that if patients were treated with ibrutinib they would have had statistically better PFS and OS outcomes. The robustness of these published analyses, therefore, provides both clinicians and payers with certainty regarding the interpretation of the comparative effectiveness of ibrutinib versus RW treatments in CLL.

PHEDRA is a unique project that extracts and utilizes patient-level RWD to identify detailed treatment patterns in the current management of hematological malignancies and to compare ibrutinib effectiveness (from both RWD and RCTs) to those of standard-of-care regimens in the clinical setting.

An important challenge when working with multisource data is finding a feasible approach for harmonizing all data, so that the analysis algorithm can be applied in a consistent manner for the same types of data across the different databases, without losing information provided by the individual source data and without changing the routine process of local data capture and collection. This is achieved by the harmonized integration of data into the CDM and the associated data handling procedures. As with other study designs producing RWD, there are some limitations related to the PHEDRA project and the interpretation of the results. Namely, the patient populations are not randomized, selected or controlled as would be the case in RCTs (though this does allow for the inclusion of a more heterogeneous population that is more typical of patients in clinical practice). Further, the geographical representativeness is very much linked to the existing database reach (i.e., a regional/single hospital database will only include patients from a restricted geographical region), although it is not assumed that these are systematically representative of the national patient population.

Selection bias is a possibility, as certain healthcare centers (such as hospitals) may only see patients at a defined stage, or at the end stage of their disease, and not necessarily from diagnosis. Hence, the availability of patients

for inclusion in the RW database might impact the end points of interest, although the direction of this impact is ambiguous. However, testing and adjusting for heterogeneity is conducted by using appropriate statistical approaches. If significant differences exist, stratification by (type of) center can be applied, for instance. Furthermore, local reimbursement criteria may exist that could influence physicians' choice (for any treatment approved at EMA level but not necessarily reimbursed/used in all countries involved), for which we are unable to adjust, and this is another limitation of real-world evidence (RWE) projects such as PHEDRA. Inherent to RWD, there is a risk of measurement error in the variables that are collected. The ways that this could compromise the precision of descriptive analyses are as follows: patient variables that are mentioned in patients' files but not reported/collected as part of the primary database structure; inconsistency of data formatting during the initial primary collection (i.e., variables collected in open text fields not usable versus data collected as a single box to tick or one item selected from a dropdown list; missing data; and lack of monitoring during the primary data collection versus RCTs that are fully controlled/validated).

One of the main end points in the PHEDRA analyses is OS, an objective end point for which the date of death is well reported. However, OS is a long-term outcome, and is impacted by subsequent line of therapy. Therefore, the relative treatment effect is likely to be most impacted when the analyses are restricted to a treatment-naïve population only. PFS, the other main end point of the PHEDRA analyses, has the advantage of not being sensitive to subsequent therapies. However, the date of the progression is often missing in RWD. In that case, for patients who initiate subsequent therapy, the date of initiation of TTNT is assigned as a proxy for progression date. As TTNT is expected to be longer than PFS (as agreed with the clinicians who own the RW databases), this approach is conservative when comparisons are made between RWD and ibrutinib clinical trial data. It should be recognized that even when the date of progression is available in the RW, it may reflect the date when the patient returned to the clinic with symptoms of progressive disease, rather than the actual time point at which disease progression occurred. This contrasts with RCTs, where a time point for assessment of disease status postremission, and criteria for determining that status, is often predefined in the protocol.

As a general rule, if data are missing for the end points, we apply a proxy that will not favor ibrutinib RCT results. For example, when data on baseline characteristics are different between the RW cohorts and the RCT, or are missing, approaches to impute the values are considered and adjustments can be made to account for the differences. If these adjustments do not affect an improvement in HR that favors ibrutinib, we can be confident that the potential residual confounding factors have not biased the results, and that the adjusted HR is a conservative estimate of the true HR.

The limitations outlined here should be carefully considered in the interpretation of the PHEDRA analyses. In all analyses presented so far, the adjusted HRs for PFS and OS across databases and analyses suggest that if these RW patients were treated with ibrutinib, they would have had statistically better outcomes. While these limitations could affect the precise estimate of the true HR of ibrutinib versus RW treatments, these limitations are unlikely to affect the direction of the HRs or the statistical significance. Furthermore, because conservative approaches are taken to minimize any potential bias, the results from PHEDRA are highly relevant and meaningful in comparing the effectiveness of ibrutinib with that of RW treatments. Many of these limitations apply to any noninterventional RWD projects [11], but do not detract from their value in complementing the data obtained from controlled studies, such as RCTs. Therefore, these limitations should not impact the direction or the significance of the results presented and these limitations should not impact the value these results demonstrate to both clinicians and payers.

Conclusion

PHEDRA is a unique project that collates, extracts and validates RWD from existing databases to evaluate treatment patterns and comparative effectiveness in CLL, MCL and WM, via adjusted comparisons with RCT data. PHEDRA is the largest RWD project to date in hematological malignancies using longitudinal patient-level data across multiple countries. The results of the PHEDRA analyses provide clinicians and payers with more certainty of the comparative effectiveness of treatments and findings are anticipated to help to improve the understanding of standards of care and should inform treatment decisions for patients with certain hematological malignancies across multiple lines of therapy. Although there are limitations in the interpretation of the resulting RWE generated by PHEDRA, these are intrinsic to any RWE analyses, and where there is uncertainty, more conservative approaches are used. Similar to any observational comparative analyses, residual confounding cannot be excluded. As RWD matures and newer medicines emerge, comparisons of innovative medicines using similar approaches will be increasingly of value.

Executive summary

- Platform for Haematology in EMEA: Data for Real World Analysis (PHEDRA) is a novel and unique, noninterventional, secondary-use real-world data (RWD) project which collaborates with existing databases in the Europe, Middle East and Africa region to gather patient-level data on chronic lymphocytic leukemia, mantle cell lymphoma and Waldenström's macroglobulinemia.
- The objectives are to understand current RW treatment patterns, as well as to compare the comparative effectiveness of ibrutinib (based on RWD and randomized clinical trial data) and alternative treatments used in RW practice.
- PHEDRA is planned to last for 8 years with regular analyses on longitudinal retrospective patient data.
- The two main end points of interest of the comparative analyses are overall survival and progression-free survival.
- PHEDRA methodology consists of data extraction, deidentification, cleaning, validation and harmonization into a Common Data Model, which enables standard statistical analyses on patient-level data for single or multiple databases.
- Limitations of the study design are intrinsic to RWD and are accounted for in the statistical considerations of the project to ensure that the most conservative measures are taken.
- The selected RW databases have collected information on a combined patient population of 2840 patients from four countries between 1990 and 2017.
- PHEDRA will inform physicians on current RW treatment practice and outcomes in order to enable optimal treatment decisions for patients.
- PHEDRA will inform health authorities and payers on the comparative effectiveness of current and new treatment regimens for chronic lymphocytic leukemia, mantle cell lymphoma and Waldenström's macroglobulinemia.

Acknowledgements

The authors thank the research collaborators for their participation and their efforts in maintaining their databases.

Financial & competing interests disclosure

The PHEDRA project is sponsored by Janssen Pharmaceutica NV and Pharmacyclics, LLC. The real-world databases utilized for PHEDRA are not owned by Janssen Pharmaceutica NV or QuintilesIMS. R Hermans, F MacDougall and D Lestelle are employees of QuintilesIMS, which is the vendor paid by Janssen Pharmaceutica NV to perform the extracting/cleaning/harmonization of the data; J Diels is an employee and equity owner of Janssen Pharmaceutica NV; H Besson, N Healy, J Garside and W Iraqi are employees of Janssen Pharmaceutica NV. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Writing assistance was provided by Jennifer Fuchs of PAREXEL and was funded by Janssen Global Services, LLC.

Ethical conduct of research

It is the responsibility of each database owner to ensure that any necessary approvals are in place for the sharing of existing deidentified patient-level data with QuintilesIMS on behalf of Janssen Pharmaceutica NV. Patient-level data are fully deidentified for analyses; reidentification was not permitted.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

- 1 Hahn OM, Schilsky RL. Randomized controlled trials and comparative effectiveness research. *J. Clin. Oncol.* 30(34), 4194–4201 (2012).
- **Review of the role of randomized clinical trials (RCTs) in comparative effectiveness research to improve the translation gap between clinical research and real-world (RW) clinical practice in oncology.**
- 2 Garrison LP Jr, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report. *Value Health* 10(5), 326–335 (2007).
- 3 Stanley K. Design of randomized controlled trials. *Circulation* 115(9), 1164–1169 (2007).

- 4 Moen F, Svensson J, Steen Carlsson K. Assessing the value of cancer treatments from real-world data – Issues, empirical examples and lessons learnt. *J. Cancer Policy* 11, 32–37 (2017).
- 5 Mahajan R. Real-world data: additional source for making clinical decisions. *Int. J. Appl. Basic Med. Res.* 5(2), 82 (2015).
- **Summarizes the importance of RW data (RWD) to support RCT data in evaluating treatment patterns and health outcomes of different treatments.**
- 6 Imbruvica. *Summary of product characteristics EU*. Janssen Pharmaceutica, High Wycombe, UK (2017).
- 7 Imbruvica. *Prescribing information US*. Janssen Biotech, Inc., Sunnyvale, CA, USA (2017).
- 8 Burger JA, Tedeschi A, Barr PM *et al.* Ibrutinib as initial therapy for patients with chronic lymphocytic leukemia. *N. Engl. J. Med.* 373(25), 2425–2437 (2015).
- **This RCT (RESONATE-2TM) demonstrates that ibrutinib was superior to chlorambucil in treatment-naïve patients with chronic lymphocytic leukemia (CLL) or small lymphocytic lymphoma (SLL), assessed by progression-free survival (PFS), overall survival (OS), response rate and improvement in hematologic variables.**
- 9 Byrd JC, Brown JR, O'Brien S *et al.* Ibrutinib versus ofatumumab in previously treated chronic lymphoid leukemia. *N. Engl. J. Med.* 371(3), 213–223 (2014).
- **This RCT (RESONATETM) in patients with relapsed or refractory CLL/SLL shows that ibrutinib compared with ofatumumab significantly improved PFS, OS and response rate.**
- 10 Chanan-Khan A, Cramer P, Demirkan F *et al.* Ibrutinib combined with bendamustine and rituximab compared with placebo, bendamustine, and rituximab for previously treated chronic lymphocytic leukaemia or small lymphocytic lymphoma (HELIOS): a randomised, double-blind, Phase III study. *Lancet Oncol.* 17(2), 200–211 (2016).
- 11 Schilsky RL. Finding the evidence in real-world evidence: moving from data to information to knowledge. *J. Am. Coll. Surg.* 224(1), 1–7 (2017).
- **Highlights the advantages and limitations inherent to RWD, and the importance of capturing the longitudinal experience of each cancer patient to guide clinical care.**
- 12 Robinson S, Brooks R, Kotiadis K, Van Der Zee DJ. *Conceptual Modeling for Discrete-Event Simulation*. Taylor & Francis Group, Boca Raton, FL, USA (2010).
- 13 Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 11(5), 550–560 (2000).
- 14 Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. *J. Am. Stat. Assoc.* 84(408), 1074 (1989).
- 15 Hernan MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am. J. Epidemiol.* 183(8), 758–764 (2016).
- **Outlines a framework for comparative effectiveness data, which helps to avoid common methodologic pitfalls, and proposes including the same patient multiple times for increased statistical efficiency.**
- 16 Byrd JC, Hillmen P, O'Brien SM *et al.* Long-term efficacy and safety with ibrutinib (ibr) in previously treated chronic lymphocytic leukemia (CLL): up to four years follow-up of the RESONATE study. Presented at: *American Society of Clinical Oncology Annual Meeting*. Chicago, IL, USA, 2–6 June (2017).
- 17 Doubek M, Obrtlíková P, Špaček M *et al.* Ibrutinib single agent vs standard of care for patients with relapsed/refractory chronic lymphocytic leukemia (CLL): an adjusted comparison of RESONATETM with the CLLEAR database. Presented at: *European Hematology Association 21st Annual Congress*. Copenhagen, Denmark, 9–12 June 2016.
- **Investigation of clinical outcomes suggesting that ibrutinib administered to patients with relapsed or refractory CLL in RESONATETM was more effective than physician's choice among individuals from an RW cohort of CLLEAR, with results consistent regardless of prior standard-of-care regimens.**
- 18 Salles G, Baseggio L, Bachy E *et al.* Single-agent ibrutinib vs standard of care for patients with treatment-naïve and relapsed/refractory CLL: an adjusted comparison of RESONATE-2TM and RESONATETM with the French Lyon-Sud database. Presented at: *American Society of Hematology Annual Meeting*. San Diego, CA, USA, 3–6 December 2016.
- **First adjusted (indirect) comparison of ibrutinib clinical trial data (RESONATE-2TM and RESONATETM trials) with RW physician's choice-treated patients with treatment-naïve or refractory CLL showing ibrutinib significantly improved both PFS and OS versus commonly used regimens.**
- 19 Doubek M, Bachy E, Špaček M *et al.* Single-agent ibrutinib versus real-world treatment for patients with treatment-naïve (TN) chronic lymphocytic leukemia (CLL): an adjusted comparison of RESONATE-2TM with the CLLEAR and Lyon-Sud databases. Presented at: *European Hematology Association 22nd Annual Congress*. Madrid, Spain, 22–25 June 2017.