Czech Registry of Monoclonal Gammopathies – Technical Solution, Data Collection and Visualisation

Český Registr monoklonálních gamapatií – technické řešení, sběr dat a jejich vizualizace

Brozova L.[1], Schwarz D.[1,2], Snabl I.[1], Kalina J.[3], Pavlickova B.[2], Komenda M.[1], Jarkovsky J.[1], Nemec P.[2], Horinek D.[4], Stefanikova Z.[5], Pour L.[6], Hajek R.[4], Maisnar V.[7]

[1] Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic
[2] Institute of Biostatistics and Analyses Ltd, Brno, Czech Republic
[3] Research Centre for Toxic Compounds in the Environment, Masaryk University, Brno, Czech Republic
[4] Department of Haematooncology, University Hospital Ostrava and the Faculty of Medicine, University of Ostrava, Czech Republic
[5] Department of Haematology and Blood Transfusion, University Hospital Bratislava, Slovak Republic
[6] Department of Internal Medicine – Hematology and Oncology, University Hospital Brno, Czech Republic
[7] 4th Department of Internal Medicine – Hematology, Charles University Hospital, Hradec Kralove, Czech Republic

Summary

Background: The Registry of Monoclonal Gammopathies (RMG) was established by the Czech Myeloma Group in 2007. RMG is a registry designed for the collection of clinical data concerning diagnosis, treatment, treatment results and survival of patients with monoclonal gammopathies. Data on patients with monoclonal gammopathy of undetermined significance (MGUS), Waldenström macroglobulinaemia (WM), multiple myeloma (MM) or primary AL ("amyloid light-chain") amyloidosis are collected in the registry. ***Data:*** Nineteen Czech centres and four Slovak centres currently contribute to the registry. The registry currently contains records on more than 5,000 patients with MM, almost 3,000 patients with MGUS, 170 patients with WM and 26 patients with primary AL amyloidosis, i.e. more than 8,000 records on patients with monoclonal gammopathies altogether. ***Results:*** This paper describes technology employed for the collection, storage and subsequent online visualisation of data. The CLADE-IS platform is introduced as a new system for the collection and storage of data from the registry. The form structure and functions of the new system are described for all diagnoses in general; these functions facilitate data entry to the registry and minimise the error rate in data. Publicly available online visualisations of data on patients with MGUS, WM, MM or primary AL amyloidosis from all Czech or Slovak centres are introduced, together with authenticated visualisations of data on patients with MM from selected centres. ***Conclusion:*** The RMG represents a data basis that makes it possible to monitor the disease course in patients with monoclonal gammopathies on the population level.

Key words

Registry of Monoclonal Gammopathies – RMG – registries – monoclonal gammopathies – CLADE-IS – data visualisation – database

This study was supported by the Ministry of Health of the Czech Republic (project identifier: DRO-FNOs/2017).

Práce byla podpořena projektem Ministerstva zdravotnictví České republiky (DRO-FNOs/2017).

The Editorial Board declares that the manuscript met the ICMJE recommendation for biomedical papers.

Redakční rada potvrzuje, že rukopis práce splnil ICMJE kritéria pro publikace zasílané do biomedicínských časopisů.

Mgr. Lucie Brozova
Institute of Biostatistics and Analyses
Faculty of Medicine
Masaryk University
Netroufalky 797/5
625 00 Brno
Czech Republic
e-mail: brozova@iba.muni.cz

Submitted/Obdrženo: 13. 6. 2017
Accepted/Přijato: 19. 6. 2017

doi: 10.14735/amko20172S43

Souhrn

Východiska: V roce 2007 byl Českou myelomovou skupinou založen Registr monoklonálních gamapatií (RMG). RMG je registr určený pro sběr klinických dat týkajících se diagnózy, léčby, jejích výsledků a přežití u pacientů s monoklonálními gamapatiemi. V registru jsou sbírána data pacientů s monoklonální gamapatií nejasného významu (monoclonal gammopathy of undetermined significance – MGUS), Waldenströmovou makroglobulinemií (WM), mnohočetným myelomem (MM) nebo primární AL („amyloid light-chain") amyloidózou. ***Data:*** V současné době do registru přispívá 19 českých a 4 slovenská centra. V registru je v současnosti evidováno více než 5 000 pacientů s MM, téměř 3 000 pacientů s MGUS, 170 pacientů s WM a 26 pacientů s primární AL amyloidózou; registr tak disponuje více než 8 000 pacienty s monoklonálními gamapatiemi. ***Výsledky:*** Článek je věnován popisu technologií využitých pro sběr, uložení dat a jejich následné online vizualizace. Představena je platforma CLADE-IS jako nový systém pro sběr a uchování dat z registru. Obecně pro všechny diagnózy je popsána struktura formulářů a funkcionality nového systému, které usnadňují zadávání nových dat do registru a minimalizují chybovost v datech. Představena je veřejně dostupná online vizualizace dat pacientů s MGUS, WM, MM nebo primární AL amyloidózou pro všechna česká nebo slovenská centra a autentizovaná vizualizace dat pacientů s MM z vybraných center. ***Závěr:*** RMG představuje datovou základnu, díky které lze monitorovat průběh onemocnění u pacientů s monoklonálními gamapatiemi na populační úrovni.

Klíčová slova

Registr monoklonálních gamapatií – RMG – registry – monoklonální gamapatie – CLADE-IS – vizualizace dat – databáze

Introduction

The Registry of Monoclonal Gammopathies (RMG) is one of the main projects of the Czech Myeloma Group (http://www.myeloma.cz). The registry is run by the Institute of Biostatistics and Analyses of the Masaryk University (http://www.iba.muni.cz) in cooperation with the Institute of Biostatistics and Analyses Ltd, a spin-off company of the Masaryk University (http://www.biostatistika.cz/), hereinafter referred to as IBA. IBA is also responsible for the technological solution of the registry. RMG is one of many clinical registries run by IBA (visit http://www.registry.cz to find out more). IBA activities include scientific research, solving research project and providing related services, particularly in the areas of analysis of biological and clinical data, organisation and management of clinical trials, development of software and ICT (information and communication technology) applications. Four IBA departments contribute to a smooth running of each registry – data management, project management, department for data analysis and IT department.

The main objectives of this paper are:

1. to introduce the RMG project,
2. to introduce the new database system as a new data basis,
3. to describe data collection, functions of the new system and structure of the registry,
4. to demonstrate and to describe interactive browsers for online visualisation of data from the registry.

Registry of Monoclonal Gammopathies

The RMG was established by the Czech Myeloma Group in 2007 [1]. The registry was established in order to collect data on patients with monoclonal gammopathies in order to monitor the disease incidence, how individual treatment modalities are used, what are their results including adverse effects, and to monitor the patients' survival in the long term. Monoclonal gammopathies are a collective name for a heterogeneous group of diseases that are characterised by the proliferation of one or more clones of differentiated B-cells producing a monoclonal immunoglobulin, which is sometimes referred to as the monoclonal protein (M-protein) or paraprotein [2]. The registry was originally focused on the collection of data on two different gammopathies – monoclonal gammopathy of undetermined significance (MGUS) and multiple myeloma (MM), i.e. a malignancy of plasma cells characterised by very diverse clinical symptoms, such as headaches caused by osteolysis, disorders of kidney function, reduced blood cell formation, frequent infections etc. The registry was extended in 2014 – forms for the collection of data on patients with Waldenström macroglobulinaemia (WM) and primary or MM-associated AL ("amyloid light-chain") amyloidosis were added.

RMG is a database suitable for parametric monitoring of patients. Cooperation of important treatment centres is key to success, because an adequate amount of representative data needs to be collected. A total of 23 treatment centres in the Czech Republic and in Slovakia currently contribute to the registry (Fig. 1). The main Czech centres involve University Hospital Brno, University Hospital Hradec Kralove, University Hospital Olomouc, University Hospital Ostrava, University Hospital Plzen, University Hospital Kralovske Vinohrady and General University Hospital in Prague. Each centre has access to its data at any time; summary analyses are based on data from those centres which have agreed to provide their data for analysis. The project is open for cooperation to all treatment centres in the Czech Republic, Slovakia and other European countries. The project has been designed as a study with both retrospective and prospective patient recruitment. All patients are asked to sign an informed consent approving the registration of their clinical data into the registry. The registry contains data on more than 8,000 patients; specifically, data on more than 5,000 patients with MM, almost 3,000 patients with MGUS, 170 patients with WM and 26 patients with primary AL amyloidosis have been recorded into the registry so far. RMG is therefore one of the largest and most comprehensive registries collecting data on patients with monoclonal gammopathies. RMG has been designed as a purely observational, research-focused and epidemiological study. Treatment with all available therapies is recorded into the registry, thus making RMG a valuable

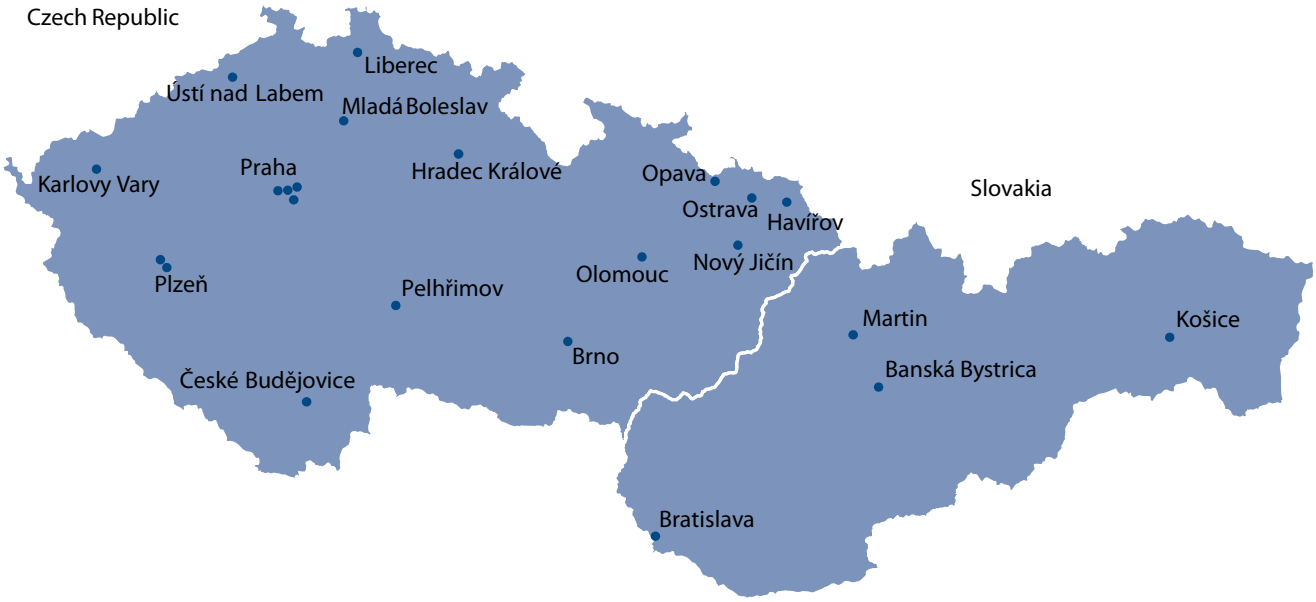

Fig. 1. Participating centres from the Czech Republic and Slovakia.

source of data from real clinical practice. Visit https://rmg.healthregistry.org to find more information on the registry itself, the list of centres with respective numbers of diagnosed patients, registry entry for authorised users, online visualisations of data, a form requesting data analysis, list of publications based on data from the registry etc.

CLADE-IS – a new online platform for data registration

The technological background of the registry has been built on the CLADE-IS platform (Clinical Data Warehousing Information System), which is software specifically designed for a repeated generation of electronic data capture (EDC) systems used in research management in clinical practice and in healthcare in general. Apart from classic randomised control trials (RCTs), CLADE-IS also supports the design and running of observational non-interventional studies, recently also referred to as "RWD/RWE" (real-world data/real-world evidence) studies. In terms of processes, the CLADE-IS platform covers all necessary steps to be made when organising a clinical trial, in accordance with the so-called good clinical practice, among others. All modules within this platform (Adminer, Designer, Reporter, Datavisor) – controlling various agendas such as the management of users and centres, design of forms with skip logic and validation rules, operating reports, data exports or graphic visualisation of data – are available online via a web browser, and therefore do not require any desktop application to be installed by the user (i.e. investigator).

In the long-term perspective, software development within the CLADE-IS platform is oriented towards open-source technologies and products (database layer – PostgreSQL, application layer – PHP/Symfony, front-end – ZURB Foundation, operation system – Linux Ubuntu LTS, deployment and virtualisation – KVM). Institute of Biostatistics and Analyses Ltd, provider of the CLADE-IS platform, prefers to work with open-source software and chooses technologies with an already existing long-term support by developer communities or even by large companies releasing their source codes as an attmpt to enhance the development dynamics. The CLADE-IS platform has replaced the formerly used system based on a proprietary technology (Microsoft, Oracle) which, although it had good results, did not meet the growing requirements anymore, namely the implementation of new functions and orientation to open-source solutions.

The CLADE-IS platform has dealt with the requirements and needs of the registry thanks to three important characteristics – 1. robustness, 2. usability, 3. openness. The robustness aspect describes the ability to define the structure of a wide range of study types – from randomised control trials (RCTs) to non-interventional observational trials or registries of qualitative data describing RWD/RWE trends. Data for RWE in clinical research cannot currently be referred to as big data because they fail to meet the definition of big data (volume, velocity, variety); nevertheless, they have many properties in common. Data variety is one of them; CLADE-IS deals with this aspect by using an original hybrid data model combining the properties of a data warehouse and a NoSQL database for unstructured data. Usability of the CLADE-IS platform is based on the application of front-end technologies that meet the requirements on responsive web design (i.e. an optimal display in a wide range of devices, from smartphones to widescreen desktop mon-

itors) and on the adherence to principles of designing the user interface as "mobile first". In this way, CLADE-IS brings breakthrough elements from the "Philosophy of UX" (User Experience) to the healthcare segment, which is often based on outdated IT solutions. This feature is important for the automation of selected agendas, such as online calculations or adverse event reporting. Openness of the system lies in its ability to communicate with other parties' systems and in the possibility of further extensions. The communication is carried out via REST API (Representational State Transfer Application Interface) and involves the possibility to obtain and to process various data from wearable technology (various biosensors or devices used in fitness and well-being) or to exchange data with international registries.

Further development of CLADE-IS is directed at improving functions in the domains of visual analytics (see below) and virtual patients, supported by unstructured data mining and by machine learning.

Structure and principles of data collection in the RMG registry

Data collection

The frequency of patients' visits to their physicians depends on their diagnosis and disease severity. At the beginning, MGUS patients have appointments every 2 or 3 months. If the patient's condition is stable, follow-up intervals increase to once every 6 to 12 months. Patients with MM, WM or AL amyloidosis visit their physicians depending on their individual condition and treatment. Patients in remission (i.e. without treatment) are usually invited to follow-up visits every 3 months. As regards patients enrolled in clinical trials, the procedure is individual, depending on the standard operating procedure (SOP) of a given trial.

Collection and updates of data in the registry are controlled by the registry coordinator. Data are entered into the registry by data managers from individual centres. In MGUS diagnosis, data in the registry are updated once a year; in WM, MM and AL amyloidosis, updates are required twice a year (each record should be updated before 30 June and before 31 December of a given year). A random check of entered data for selected patients is performed once a year by the registry coordinator, who compares these data from the registry with medical documentation in the participating centre. Apart from these tasks, the registry coordinator arranges training of new registry users as well as current users (aiming to inform them of new recommendations by CMG on data collection), helping users with entering data into the registry, having new user accounts created or existing user accounts deleted, etc.

Important functions of the system

Data originally stored in a former database system (based on a modified version of the TrialDB system) were converted to a new CLADE-IS system in 2017. Each individual user can use his/her username and password to enter the registry at https://rmg.data-warehouse.cz/login. Users have different rights according to their roles – for example, representatives of a given treatment centre can only see data from that centre. Data in the registry are de-identified, i.e. each patient in the registry is represented by an identification code that makes it impossible to reveal the patient's identity. Entered data can be exported from the registry to a table in Excel format. The IBA helpdesk sends regular monthly reports to registry users, summing up the numbers of patients with individual diagnoses recorded by the respective centre.

The new CLADE-IS offers important functions that not only make data entry into the registry easier and faster for users, but also minimise the error rate in data:

1. The system enables users to search for patients according to defined characteristics (for example, when the user needs to find patients with a specific type of treatment in order update their data).
2. Skip logic rules contribute to shortening of otherwise lengthy forms and require specific items to be filled in. For example, if the user indicates that a cytogenetic examination was carried out, the system displays individual items (presence of translocations, gains etc.) and requires them to be filled in. By contrast, if the user indicates that the cytogenetic examination was not carried out, the related items remain hidden. In this way, the form length changes dynamically.
3. Validation rules are set in some important parts of individual forms, making it impossible for the user to enter wrong data into the registry (for example, a patient's treatment cannot be terminated before it is started). Limits are set in selected parameters, requiring the entered value to fall between them (for example, percentages are limited to a minimum of 0 and a maximum of 100).
4. All items in the registry can be classified into two categories – mandatory and optional. If an item is marked as mandatory, the expert group (CMG) requires it to be filled in. Not filling such an item will prevent saving the form as "valid", which is an indication for the user that not all important items have been filled in. If the user considers the form to be completely filled in, he/she saves it as "completed". The system will subsequently evaluate whether all mandatory items are filled in; if so, the state of the form is changed to "valid"; if not, the system will prompt the user to fill the mandatory items (the text "this field is required" will be displayed in a red box under the item title). In situations when the user does not know the value of a mandatory item, the form must be saved as "pending". Some parameters can be marked as non-evaluable (for example, this can be used in a situation where a given treatment centre does not perform a specific test to find the value of a certain parameter); after marking the mandatory item as non-evaluable, the form can be saved as "completed".

Structure of forms

When working with the RMG database, users create and fill in forms for four groups of diagnoses – MGUS, WM, MM and primary AL amyloidosis. In general, three types of forms are filled in for each of these diagnoses:

1. **Diagnostics** – Characteristics of the patient's condition at the time of diagnosis. The evaluated parameters usually involve – date of diagnosis, characteristics describing the M-protein, performance status, disease stage, cytological and histological examination of bone marrow, evaluation of presence of bone lesions and extramedullary mass, results of biochemical, flow-cytometric and cytogenetic examinations.
2. **Follow-up/Treatment** – Description of a patient's follow-up over time. In MGUS patients, one form is repeatedly used to record paraprotein levels over time; date of evaluation, level of M protein and potential comorbidities are recorded. In WM, MM or primary AL amyloidosis patients, the form is created repeatedly for each treatment line. In general, each Treatment form describes – line of treatment, clinical symptoms of the disease, patient's characteristics at the time of treatment initiation (e.g. biochemistry), dates of treatment initiation and treatment termination, description of the treatment itself (drugs, dosage, whether a transplantation was carried out etc.), assessment of treatment effectiveness (treatment response), adverse effects/treatment toxicity and date of progression after treatment. One treatment switch can be recorded into the form; a switch should be mostly performed due to toxicity of the previous therapy. If the treatment is not successful, change of treatment regimen should always lead to the creation of another form for a new treatment line.
3. **Current status** – Description of a patient's condition at the last follow-up visit. In MGUS patients, progression to malignancy is evaluated; in WM patients, transformation to another disease and the presence of another malignancy is evaluated. The patient's condition as of the date of the last visit is recorded for all diagnoses; if the patient was alive, it is recorded whether treatment was administered; if the patient died, date and cause of death are recorded.

The content of forms is always adapted to the diagnosis of interest; for example, the diagnostic form for primary AL amyloidosis also includes the biopsy of amyloid. The diagnostic form for MM is also created for patients diagnosed with smouldering MM (SMM). The user then ticks off in the form whether the patient meets the criteria for SMM diagnosis. If MM and AL amyloidosis were diagnosed in a patient at the same time, the user should also fill in a diagnostic form for secondary MM-associated AL amyloidosis (apart from forms related to MM).

Online visualisation of data

With the growing amount of information available across the Internet, inevitably there is a growing need not only to process and to store data, but also to present data in a comprehensible manner. The latter issue is addressed by visual analytics, a dynamically developing field of data analysis and visualisation. Using advanced statistical methods, available data are searched for interesting and at first glance invisible features and relations, and these are subsequently presented to the target audience in the form of a summary report. The domain of data visualisation is characterised by a wide range of possibilities arising from modern technologies that are used in the development and subsequent implementation of web applications. Online visualisations provide a well-arranged view of required information, together with interactive elements and animated effects. A navigation panel for the specification of data selection and for setting the view itself (in the form of data tables, graphs or maps) usually forms an integral part of such visualisations. As part of the RMG project, an interactive browser has been implemented over data that had been collected via the CLADE-IS platform. With regard to the nature of published information, the visualisation is divided into a public part (Fig. 2 – available online at www.linkos.cz) and an authenticated part (Fig. 3A–C – available online at www.linkos.cz), revealing only relevant information to a given user, according to his/her level of access rights. Access to both types of visualisation is available at the RMG project website, namely in the section "Interactive browser" (https://rmg.health-registry.org/index.php?pg=data). All outputs are automatically processed, solely based on data contained in the registry; validity of these data is controlled only by the skip logic and validation rules, both of which had been previously set in the system.

Technology

The RMG project website, including the interactive browser, has been developed using the following technologies:

- the LAMP platform for the implementation of web applications (Linux, Apache, MySQL, PHP),
- scripting languages JavaScript and R (support for the OpenCPU system) to process and to present data online,
- WKHTML component to convert the outputs into PDF format in the authenticated part of visualisations.

Our own library of visualisation components has been used to develop the individual graphical components. This library loads data in the JSON format (JavaScript Object Notation) and uses the d3.js component to display the graphical representation in the SVG format (Scalable Vector Graphics). Fonts, colours and other design attributes are defined via CSS 3 (Cascading Style Sheets). This technology has been used to compose pie charts and bar charts into the interactive browser. Furthermore, the authenticated version of data visualisation is supplemented with survival analysis, which is presented via Kaplan-Meier curves and their descriptive statistics. In order to calculate a survival curve, the application sends a query to the R package (described in the following section), which is made accessible by the OpenCPU system. The result, which is received in the form of JSON data, is subsequently processed and displayed online.

RMG package

In order to perform calculations related to survival analysis, the integrated development environment RStudio 1.0.136 was used to develop and to implement an R package which employs the non-parametric Kaplan-Meier

estimate to evaluate the probability of event-free survival in the following four time-to-event endpoints – overall survival (OS), time to progression (TTP), progression-free survival (PFS) and duration of response (DOR). Data processing for visualisation is performed by a script running on an OpenCPU server, which provides the HTTP API (Hypertext Transfer Protocol Application Programming Interface) to calculate survival analysis, receives input data via the POST method and returns results both in the form of source data for graph display and in the form of statistics. The package itself contains a single function entitled rmg(), which is called by a specific address at the OpenCPU server (implemented via a OpenCPU package in the R language) and which employs the POST method to receive an array of eight vectors (in the JSON format) representing the values of parameters needed to calculate the time-to-event endpoints in individual patients.

The rmg() function calculates the time-to-event endpoints in accordance with IMWG (International Myeloma Working Group) definitions [3]. In event-free patients, the interval length is calculated as of the date of the last follow-up and these patients are censored in the analysis. Patients with missing values of parameters needed for classification into groups (such as age or treatment line) or with negative values of time to event (as a result of incorrectly entered data) are excluded from the analysis.

The RMG package uses previously published packages survival, splines and RColorBrewer, which are publicly available via the CRAN (Comprehensive R Archive Network) repository. In order to calculate the Kaplan-Meier estimate, the following functions from the survival package are used: 1. the Surv() function, which creates an object of class "survival"; 2. the survfit() function, which considers this object as a dependent variable to define the survival model as a background for the calculation of survival curves. The graphs contain a single curve if survival is evaluated together for all patients in a given selection; or more curves if patients are classified into groups (such as two curves for two groups of patients, e.g. those with transplantation versus those without transplantation in a given line of treatment). If more curves are presented in a graph, the statistical significance of the difference between the curves is quantified via the p-value of log-rank test; this value is obtained from the survdiff() function. This procedure is repeated four times, separately for each time-to-event endpoint.

This process results in a graphical output for each time-to-event endpoint, in the form of one or more (differently coloured) curves drawn over a time interval from 0 to 96 months. Each graph is accompanied by a table, which describes the course of individual curves in more detail. Specifically, the table provides the overall number of patients, the overall number of events, median survival for a given time interval supplemented with 95% CI, and the probability of event-free survival in 12, 24, 36, 60 and 120 months, again supplemented with 95% CI.

Contents and work with browsers

From the user's point of view, work with the browser can be described as follows – 1. starting the application in a web browser; 2. in authenticated version, the user's identity is verified and predefined user rights are assigned; 3. data are selected using a set of filters; 4. visualisations are displayed. Updates of data to be presented are performed once a day, in nighttime hours of CET (Central European Time). Apart from several views of data on the webpage, the authenticated version also offers a full export of all available visualisations in the PDF format. Reports creation and their conversion to PDF format is done by a service on a dedicated server; due to high computational demands, requests on exports are lined up in a queue, which is subsequently processed by the service. Downloading the PDF export and putting it into the repository is the last step. The initial page shows the list of reports, with ten most recent ones being archived.

Figure 2 (online) demonstrates the web environment of a publicly available visualisation. In terms of contents, the publicly available browser describes only basic data on patients with monoclonal gammopathies (MGUS, WM separately for asymptomatic and symptomatic forms, MM, primary AL amyloidosis). Filters in publicly available visualisations involve the type of diagnosis and the selection of treatment centres (all centres/all Czech centres/all Slovak centres). Taking into account these filters, the browser displays:

1. Basic overview – numbers of newly diagnosed patients in individual years, by sex and in total; the total number of men and women over the entire monitored period. For the MM diagnosis, these results are supplemented with the total number of patients who had met the criteria for SMM at the time of diagnosis, and the total number of patients who developed MM-associated AL amyloidosis during the follow-up period. For the primary AL amyloidosis, the results are supplemented with the total number of patients with a given type of amyloidosis (either systemic or localised).
2. Follow-up – duration of follow-up from the date of selected diagnosis to either the last evaluation or the patient's death.
3. Age structure – age structure by sex and in total.
4. Paraproteins – type of heavy chain in the paraprotein.
5. Clinical stages – ISS (International Staging System) [4] and Durie Salmon staging system [5], only available for the MM diagnosis.
6. Risk score – in patients with MGUS or MM, the proportion of groups of patients with various numbers of risk factors according to the Mayo stratification system [6] is presented.

Unlike the publicly available visualisation, the authenticated version of the browser provides more detailed overviews focused on the MM diagnosis. Access to these visualisations has been granted to representatives of all treatment centres, each of whom is entitled to display data from his/her centre only, and to top representatives of CMG, who have extended rights and are entitled to display data from any individual

centre or overall data for either Czech or Slovak centres. Figures 3A–C (online) demonstrate the environment in this part of visualisations. After login, the user can set various filters for the analysis, i.e. selects criteria to be met by patients involved in the analysis. These criteria include:

1. Centre – a centre (or a group of centres) can be selected only in cases when the user is entitled to display data from more centres.
2. Year of treatment initiation – a scroll bar can be used to select the range of years; for example, all treatments initiated in the period 2007–2010. It is important to keep in mind that treatments initiated in 2006, for example, which carry on in the period 2007–2010, would not be involved in such analysis.
3. Age (treatment initiation) – those aged ≤ 70 years or those aged > 70 years.
4. Line of treatment – 1st, 2nd, 3rd, 4th, 5th and higher lines of treatment. Each patient is presented in a data set as many times as the number of treatment lines he/she initiated. The visualisation makes it possible to select more treatment lines.
5. Clinical trial – selection of treatments that were – or were not – administered as part of clinical trials.
6. Autologous stem cell transplantation (ASCT) – selection of induction regimens which were – or were not – followed by ASCT.
7. Treatment regimen – this option currently makes it possible to select treatments in which a new therapy has been used – bortezomib, thalidomide, lenalidomide, carfilzomib, pomalidomide, ixazomib or another therapy, which involves all therapies other than the above-mentioned ones. More therapies can be selected in the filter. If a treatment combination involved bortezomib and thalidomide, for example, this combination would be included in the analysis both in the situation where only bortezomib was selected and also in the situation where only thalidomide was selected. The selection of drugs is going to be expanded together with the development of new drugs.

By default, all parameters are set to "total", indicating that the user selects all data from the centre(s) of choice. The user can also tick off the "generate PDF report" option, if he/she wants the results to be exported to a PDF file; after clicking on the "display" button, the user is redirected on the main visualisation page. The upper part of this page shows the list of ten most recent reports – if the item "state" is marked as "finished", the report is ready for download. The blue box on the main page sums up all previously set filters for patient selection as well as the overall number of patients and treatment lines involved in the evaluation. An interactive list of items is displayed below the summary of predefined filters; results are displayed after clicking on a specific item.

The items are divided into five sections:

1. **Summary lists** – this section involves the total numbers of newly diagnosed patients in individual years, numbers of newly initiated treatments in years depending on the line of treatment, treatment regimen and in total; and the year of last update of records in the registry for patients who are alive.
2. **Basic overview** – patients – basic description of patients diagnosed with MM. Visualisations are available for sex, age, duration of follow-up, ISS and performance status. Although more than one treatment line for the same patient might meet inclusion criteria, any given patient is only involved once for the evaluation in this section.
3. **Basic overview** – treatment lines – basic description of patients at the time of treatment initiation with a given treatment line; this means that the patient is involved in the evaluation as many times as the number of his/her treatment lines which met the inclusion criteria. The same characteristics are evaluated as in Section 2; furthermore, the evaluation also involves the treatment line, treatment regimen, whether ASCT was performed and whether clinical trial was involved.
4. **Final treatment response and toxicity** – this section evaluates the success of treatment by the means of treatment response (defined by the current IMWG criteria [3]) and treatment toxicity (for example, anaemia or infections).
5. **Patient survival Kaplan-Meier curves** are used to present the results of treatment intervals OS, TTP, PFS and DOR. The first part displays the survival for all treatment lines that meet inclusion criteria. In other parts, treatment lines are classified to multiple curves, according to patient characteristics (for example, ISS stage I–III, cytogenetical examination positive/negative).

All outputs available in the web environment are also included in the PDF report, where each output is presented on a separate slide. By clicking on "log out", the user leaves the visualisation and returns to the RMG project website.

Conclusion

The RMG was established in 2007. Data originally stored in a former database system (based on a modified version of the TrialDB system) were converted to a new CLADE-IS system in 2017. Revision of form structure and contents was performed together with database conversion. New validation criteria were set in order to minimise the error rate in data. Data are recorded into the system by data managers, who are supervised by the registry coordinator. Data from the registry are evaluated after the approval by representatives of participating centres. A basic overview of data on individual diagnoses from all Czech or Slovak centres together is publicly available on the registry website. A detailed online visualisation of data on MM patients from selected centres, which can be also exported to a PDF file, is available for representatives of individual centres as well as for top representatives of CMG.

RMG represents an international database designed for the collection of data on patients with monoclonal gammopathies. Nineteen Czech centres and four Slovak centres currently contribute to the registry. The registry already contains data on more than 8,000 patients with monoclonal gammopathies. RMG is therefore one of the largest registries with

regular monitoring in Europe. Data from RMG are utilised in the preparation of national guidelines for treatment of monoclonal gammopathies, serve as a basis for negotiations with the State Institute for Drug Control (SUKL), and are also used for subsequent negotiations on reimbursements of new drugs with healthcare payers. The RMG represents a data basis that makes it possible to monitor the disease course in patients with monoclonal gammopathies on the population level.

Conflicts of interest

D. S., B. P. and P. N. declare an employment relationship with the Institute of Biostatistics and Analyses Ltd. – the owner and operator of the CLADE-IS system. R. H. has a consultant or advisory relationship and received honoraria from Amgen, Bristol-Myers Squibb, Takeda, Celgene and Janssen-Cilag; conducted a clinical research project(s) funded by Takeda, Novartis, Amgen and Janssen-Cilag. V. M. consulted for Amgen, Bristol-Myers Squibb, Celgene, Janssen-Cilag and Takeda; received grant support from The Binding Site, honoraria from Amgen, Bristol-Myers Squibb, Celgene and Janssen-Cilag, and has been involved in advisory boards for Amgen, Bristol-Myers Squibb, Celgene, Janssen-Cilag and Takeda. Other authors declare they have no conflicts of interest concerning drugs and other medicinal products used in the study.

Střet zájmů

D. S., B. P. a P. N. deklarují pracovněprávní vztah ke společnosti Institut biostatistiky a analýz s.r.o. – vlastníkovi a provozovateli systému CLADE-IS. R. H. je konzultantem nebo poradcem u společností Amgen, Bristol-Myers Squibb, Takeda, Celgene a Janssen-Cilag, od kterých obdržel honoráře; provádí klinické výzkumné projekty financované společnostmi Takeda, Novartis, Amgen a Janssen-Cilag. V. M. je konzultantem u společností Amgen, Bristol-Myers Squibb, Celgene, Janssen-Cilag a Takeda; přijal grantovou podporu od The Binding Site a honoráře od Amgen, Bristol-Myers Squibb, Celgene and Janssen-Cilag; je členem poradních výborů u společností Amgen, Bristol-Myers Squibb, Celgene, Janssen-Cilag a Takeda. Ostatní autoři deklarují, že v souvislosti s předmětem studie ve smyslu léků a jiných léčivých přípravků nemají střet zájmů.

References

1. Maisnar V, Pelcová J, Klimeš D et al. RMG – Registr Monoklonálních Gamapatií. Onkologie 2011; 5(3): 138–140.
2. Kyle RA. The monoclonal gammopathies. Clin Chem 1994; 40(11 Pt 2): 2154–2161.
3. Rajkumar SV, Harousseau JL, Durie B et al. Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1. Blood 2011; 117(18): 4691–4695. doi: 10.1182/blood-2010-10-299487.
4. Greipp PR, Miguel JS, Durie BG et al. International Staging System for Multiple Myeloma. J Clin Oncol 2005; 23(15): 3412–3420.
5. Durie BG, Salmon SE. A clinical staging system for multiple myeloma. Correlation of measured myeloma cell mass with presenting clinical features, response to treatment, and survival. Cancer 1975; 36(3): 842–854.
6. Kyle RA, Durie BG, Rajkumar SV et al. Monoclonal gammopathy of undetermined significance (MGUS) and smoldering (asymptomatic) multiple myeloma: IMWG consensus perspectives risk factors for progression and guidelines for monitoring and management. Leukemia 2010; 24(6): 1121–1127. doi: 10.1038/leu.2010.60.