Monitoring Dialysis Outcomes across the World – The MONDO Global Database Consortium

Gero D. von Gersdorff a Len Usvyat b, c Daniele Marcelli d Aileen Grassmann d Cristina Marelli e Michael Etter f Jeroen P. Kooman g Albert Power h Ted Toffelmire i Yosef S. Haviv j Adrian Guinsburg e Claudia Barth k Mathias Schaller a Inga Bayh d Laura Scatizzi d Adam Tashman b Stephan Thijsse nathan W. Levin b Frank M. van der Sande g Charles Pusey h Yuedong Wang l Peter Kotanko b

a Cologne University Medical Center, Cologne, Germany; b Renal Research Institute, New York, N.Y., c Fresenius Medical Care North America, Waltham, Mass., USA; d Fresenius Medical Care, Bad Homburg, Germany; e Fresenius Medical Care Latin America, Buenos Aires, Argentina; f Fresenius Medical Care Asia-Pacific, Hong Kong, SAR, China; g Maastricht University Medical Center, Maastricht, The Netherlands; h Department of Medicine, Imperial College London, London, UK; i Queen’s University, Kingston, Ont., and Fresenius Medical Care, Richmond Hill, Ont., Canada; j Hadassah-Hebrew University Medical Center, Jerusalem, Israel; k Kuratorium für Dialyse und Nierentransplantation e.V., Neu-Isenburg, Germany; l University of California – Santa Barbara, Santa Barbara, Calif., USA

Key Words
Dialysis · Global database · Longitudinal trends · MONDO · Observational study · Registry methods

Abstract
Background/Aims: Dialysis providers frequently collect detailed longitudinal and standardized patient data, providing valuable registries of routine care. However, even large organizations are restricted to certain regions, limiting their ability to separate effects of local practice from the pathophysiology shared by most dialysis patients. To overcome this limitation, the MONDO (MONitoring Dialysis Outcomes) research consortium has created a platform for the joint analysis of data from almost 200,000 dialysis patients worldwide. Methods: We examined design and operation of MONDO as well as its methodology with respect to patient inclusion, descriptive data and other study parameters. Results: MONDO partners contribute primary databases of anonymized patient data and collaboratively analyze populations across national and regional boundaries. To that end, datasets from different electronic health record systems are converted into a uniform structure. Patients are enrolled without systematic exclusions into open cohorts representing the diversity of patients. A large number of patient level treatment and outcome data is recorded frequently and can be analyzed with little delay. Detailed variable definitions are used to determine if a parameter can be studied in a subset or all databases. Conclusion: MONDO has created a large repository of validated dialysis data, expanding the opportunities for outcome studies in dialysis patients. The density of longitudinal information facilitates in particular trend analysis. Limitations include the paucity of uniform definitions and standards regarding descriptive information (e.g. co-morbidities), which limits the identification of patient subsets. Through its global outreach, depth, breadth and size, MONDO advances the observational study of dialysis patients and care.

© 2013 S. Karger AG, Basel
**Introduction**

Dialysis providers typically monitor the treatment of their patients closely as part of their medical quality management systems. Data are collected in a standardized fashion, and care is taken to gather clinical, treatment-related and biochemical data that allow comparison between dialysis units. Provider databases are characterized by rich longitudinal follow-up information [1]. However, even large organizations are restricted to certain regions, limiting their ability to separate effects of local practice from the pathophysiology shared by most dialysis patients. One way to overcome these constraints and to study large numbers of patients is to combine databases from multiple providers from multiple countries. Observational studies of registry data face major challenges, including unrecognized sources of bias in the inclusion of subjects as well as bias introduced by data collection processes and poor data validation. In addition, the integration of multiple datasets from different electronic medical record (EMR) systems into a single repository of validated data is challenging.

Starting from an ad hoc collaboration to address a specific scientific question [2], the MONDO (MONitoring Dialysis Outcomes) initiative was founded and subsequently developed into a consortium in which a variety of academic and nonacademic institutions from around the world work together on research projects to analyze primary clinical databases of dialysis patients [3].

**Aims**

In order to assist in the interpretation of forthcoming reports from the MONDO initiative, we provide here an analysis and description of the registry protocol, focusing on aspects that may affect the results obtained from MONDO data [4, 5]. We will also highlight the goals, unique features and limitations of this data source.

**Methods**

MONDO can be viewed as a global 'registry' composed of datasets from several primary databases (fig. 1). Several components of this registry have special relevance for the interpretation of results obtained [4–6]. These relate to the organizational structure of the registry as well as to the study subjects and the data collected. Focal points of the analysis of the design and organizational structure of MONDO were the purpose and scope of this registry and its characteristics. With regard to subjects, we analyzed inclusion and exclusion policies, the population available for study and measures taken to protect patient confidentiality. Important features of data management were the methods employed to achieve a uniform and systematic way of data collection, the specificity and consistency of data definitions between databases and the approach to editing and verifying data, including the handling of missing data.

To obtain this information, a questionnaire was developed [4, 5] and sent to all directors and managers of the MONDO partner databases. This effort was supplemented by telephone interviews and discussions at several committee meetings. The general structure as well as some aspects of the MONDO initiative have already been described [3]. In addition, the structure and content of some of the primary electronic databases that collaborate in the MONDO initiative have been reported previously [1, 7–10].

**Results**

**Purpose and Structure of the Consortium**

The MONDO consortium is a research collaboration that explores established clinical databases of dialysis patients across national, regional and provider borders. It constitutes a framework for a variety of observational study designs. MONDO encompasses patients of diverse ethnic backgrounds who are treated in different medical systems and under different reimbursement policies. To facilitate joint analyses across de-identified datasets from different consortium members, primary datasets are converted into a uniform data structure for MONDO (fig. 2). The resultant MONDO analytical files, which together constitute a virtual MONDO ‘data bank’, are either kept locally for analysis if requested by the provider or analyzed centrally at the Renal Research Institute (RRI). The first converted datasets covered patients treated between January 2000 and December 2010 and updates are scheduled at regular intervals. Currently, the following providers have converted datasets for MONDO: the RRI (USA), Fresenius Medical Care (FMC) Europe, Latin America and Asia-Pacific, the Kuratorium für Dialyse und Nierentransplantation e.V. (Curatorium for Dialysis and Kidney Transplantation; KfH, Germany), Maastricht University Medical Center (The Netherlands), Imperial College London (UK) and the Hadasah-Hebrew University Medical Center (Israel). Datasets from additional partners in the Middle East, Africa, Latin America and Asia are in preparation, and the consortium is open to new members to further increase its scope. Research projects can be submitted by all consortium members as well as outside parties. Data can be analyzed in two modalities, centrally or locally, in order to integrate provider-specific regulations and demands. Common standards for analysis are followed in both cases and research papers are published jointly. Funding for the primary databases including data conversion into the
MONDO data structure is provided by each consortium member. Expenses related to data analysis and statistical support, if not performed by the members themselves, are covered by the RRI. Participation in joint meetings and other administrative costs are the responsibility of each member.

Subjects and Ethical Standards, Patient Confidentiality and Data Protection

Subjects
All participating organizations offer enrollment into a primary database to all chronic ambulatory (outpatient) dialysis patients. From a study perspective, patients are enrolled without systematic exclusions into open cohorts from the time they start receiving chronic dialysis treatment in each organization and contribute data for as long as they receive treatment there. In all participating databases, enrollment is above 90% of all patients treated. Since individual patients can only receive treatment from one organization at a time they are unique in the MONDO database. While all organizations offer standard hemodialysis and peritoneal dialysis, most also offer a variety of alternative dialysis modalities, such as nocturnal in-center, home or frequent hemodialysis. The spectrum of patients included is therefore very broad and reflects daily practice including rare events and patients that are treated infrequently in a unit or have uncommon clinical presentations. Although providers often treat a large share of the dialysis population in a respective country or region, samples in the MONDO database cannot necessarily be viewed as representative of that population. The MONDO database currently contains almost 200,000 chronic dialysis patients from 37 countries on 5 continents: Europe and Middle East: 23 countries (19 FMC, 1 KfH, 1 Imperial College, 1 Maastricht and 1 Hadassah Medical Center); Latin America: 5 countries (FMC); Asia-Pacific: 8 countries including Australia and New Zealand (FMC), and North America: 1 country (RRI; fig. 1). Overall patients’ average age is 61.8 years; there are 41% women, and 28% are diabetic. There are approximately 43 million hemodialysis and 3 million peritoneal dialysis treatments in the database along with associated clinical and laboratory parameters. Longitudinal data are also available from over 1,500 pediatric patients aged 17 years and younger.

Ethical Standards, Patient Confidentiality and Data Protection
Patient participation follows the standards of the Declaration of Helsinki and Good Epidemiological Practice.
in all organizations. All organizations are responsible for the primary collection and safeguarding of patient data in accordance with all local data protection laws and privacy protection regulations. They are also responsible for ensuring compliance with laws and regulations regarding the secondary use of data in the context of MONDO. In most instances, the anonymized patient data are analyzed by RRI. KfH does not share data but uses jointly developed statistical analysis codes and analyzes data locally in accordance with provider rules, their patients’ written consent and local privacy protection laws.

Data Collection, Validity and Comparability between Databases
Data Flow and Processing
All participating organizations use EMR to document patient care. EMR information is aggregated into a primary database (fig. 2). A data extract for MONDO is made by transforming de-identified primary data into a uniform MONDO data structure. Basic data cleaning is performed (e.g. range checks and conversion to standard units). The resultant MONDO analytical files can be analyzed centrally (a) or locally (b; e.g. KfH) using the statistical program codes developed jointly by a statistical work group.

Type of Data Collected
Chronic dialysis care is highly standardized throughout the world. The purpose of the primary databases is to monitor patient care and to identify potential for im-

---

**Fig. 2.** Data processing for MONDO. 1st level: EMR data are aggregated into a primary provider database. The processes to ensure compliance with data protection regulations differ between countries and between databases. 2nd level: sets of de-identified primary data for MONDO are identified and converted into a uniform data structure. Basic data cleaning is performed (e.g. range checks and conversion to standard units). The resultant MONDO analytical files can be analyzed centrally (a) or locally (b; e.g. KfH) using the statistical program codes developed jointly by a statistical work group.
In order to promote data completeness, some parameters for quality assurance and reimbursement issues. In the process of creating the MONDO data bank, care is taken to record the expected parameter characteristics (e.g. frequency, timing and method) for each database. Key parameters and their characteristics are summarized in table 1. Since differing parameter definitions may influence the validity of results, the suitability of data from a database is individually assessed for each MONDO project. Some characteristics and their potential effects are notable.

### Frequency of Laboratory and Treatment Parameters
In the majority of primary databases, parameters are collected at the same frequency as they are obtained from the patient (e.g. weights for each treatment and monthly to quarterly laboratory investigations). Only three databases (FMC Latin America and Asia-Pacific, and Imperial College) extract a representative monthly dataset from the EMR. Because of this abundance of longitudinal data, the effects of differences in the collection schedules between databases are often negligible.

### Timing of Laboratory Parameters
In the majority of databases, blood sampling is performed midweekly (for a typical thrice-weekly schedule), but others collect values after the 'long interval'. Although this may impact on the absolute value of a parameter, an analysis of longitudinal trends for this parameter would not be affected.

### Laboratory Assays
Laboratory results in MONDO reflect the variety of assays found in clinical practice throughout the world. The values for most of these are internationaly comparable [11], but attention was paid to convert values into consistent units (typically SI). When there were predictable differences between assays, interassay conversion was attempted whenever feasible. Thus, values for albumin were converted from the bromocresol purple to the bromocresol green method by published formulae [12] before they were deposited in the MONDO data bank.

### Case-Mix Variables
Cause of kidney failure (primary renal disease), type and number of comorbid conditions and cause of death are recorded in all data capture systems. Most EMR systems require selection of relevant diagnoses either from a code system such as ICD-9 or ICD-10, or from proprietary catalogues ('tick when present') [13, 14]. In order to promote data completeness, some EMR systems demand an entry before data of a patient can be saved. Selection and documentation of these diagnoses therefore reflect clinical assessment of the patient by the medical staff in the context of dialysis care and not codable or reimbursable conditions. Diagnoses from outside the dialysis system, such as from hospitals or other medical specialties who treat the same patient, are typically not obtained. Information will thus be more com-

---

**Specificity and Consistency of Data Definitions between Databases**

Although data collection within primary databases is uniform and consistent, data definitions between databases may differ. It is noteworthy, however, that all databases were strongly influenced by the Kidney Disease Outcomes Quality Initiative and Kidney Disease Improving Global Outcomes guideline and therefore followed similar principles in the methods and detail of collection of core monitoring parameters. Provider-specific differences also reflect country-specific legal demands on parameters for quality assurance and reimbursement issues. In the process of creating the MONDO data bank, care was taken to record the expected parameter characteristics (e.g. frequency, timing and method) for each database. Key parameters and their characteristics are summarized in table 1. Since differing parameter definitions may influence the validity of results, the suitability of data from a database is individually assessed for each MONDO project. Some characteristics and their potential effects are notable.

**Frequency of Laboratory and Treatment Parameters.** In the majority of primary databases, parameters are collected at the same frequency as they are obtained from the patient (e.g. weights for each treatment and monthly to quarterly laboratory investigations). Only three databases (FMC Latin America and Asia-Pacific, and Imperial College) extract a representative monthly dataset from the EMR. Because of this abundance of longitudinal data, the effects of differences in the collection schedules between databases are often negligible.

**Timing of Laboratory Parameters.** In the majority of databases, blood sampling is performed midweekly (for a typical thrice-weekly schedule), but others collect values after the 'long interval'. Although this may impact on the absolute value of a parameter, an analysis of longitudinal trends for this parameter would not be affected.

**Laboratory Assays.** Laboratory results in MONDO reflect the variety of assays found in clinical practice throughout the world. The values for most of these are internationally comparable [11], but attention was paid to convert values into consistent units (typically SI). When there were predictable differences between assays, interassay conversion was attempted whenever feasible. Thus, values for albumin were converted from the bromocresol purple to the bromocresol green method by published formulae [12] before they were deposited in the MONDO data bank.

**Case-Mix Variables.** Cause of kidney failure (primary renal disease), type and number of comorbid conditions and cause of death are recorded in all data capture systems. Most EMR systems require selection of relevant diagnoses either from a code system such as ICD-9 or ICD-10, or from proprietary catalogues ('tick when present') [13, 14]. In order to promote data completeness, some EMR systems demand an entry before data of a patient can be saved. Selection and documentation of these diagnoses therefore reflect clinical assessment of the patient by the medical staff in the context of dialysis care and not codable or reimbursable conditions. Diagnoses from outside the dialysis system, such as from hospitals or other medical specialties who treat the same patient, are typically not obtained. Information will thus be more com-
Table 1. Comparison of parameter characteristics between databases: key parameters that are available in all databases are shown with their frequency of collection and timing relative to a typical thrice-weekly dialysis schedule (a great number of additional parameters with standardized collections are included in MONDO from different databases)

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Background and demographics</strong></td>
<td>Age, gender, race</td>
<td>RRI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td><strong>Admission and discharge data</strong></td>
<td>Data related to admission, discharge, RRT initiation, reasons for discharge</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Medical history</strong></td>
<td>Primary renal disease, comorbid conditions, cause of death (code system)</td>
<td>ICD-9</td>
</tr>
<tr>
<td><strong>Hospitalizations</strong></td>
<td>Admission and discharge data</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Hospital discharge diagnosis (code system)</td>
<td>ICD-9</td>
</tr>
<tr>
<td><strong>Dialysis prescription</strong></td>
<td>Treatment time (per treatment)</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>eKt/V, URR (1–3 m)</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Vascular access</strong></td>
<td>Vascular access, type and dates (1–3 m)</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Vital signs</strong></td>
<td>Blood pressure (per treatment)</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Weight/volume management</strong></td>
<td>Pre- and postdialysis weight (per treatment)</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Laboratory</strong></td>
<td>Timing of blood sampling (relative to 3×/week schedule)</td>
<td>mid-week</td>
</tr>
<tr>
<td></td>
<td>Anemia management</td>
<td>Hemoglobin (m)</td>
</tr>
<tr>
<td></td>
<td>Ferritin/TSAT (1–4 m)</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Mineral and bone disease</td>
<td>Calcium, phosphorus (m)</td>
</tr>
<tr>
<td></td>
<td>PTH (every 3–6 months)</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Nutrition/inflammation</td>
<td>C-reactive protein</td>
</tr>
<tr>
<td></td>
<td>Albumin (1–4 m)</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>WBC, neutrophils, lymphocytes (1–3 m)</td>
<td>✓</td>
</tr>
</tbody>
</table>

eKt/V = equilibrated Kt/V (urea); m = monthly; n/a = not available; PTH = parathyroid hormone; RRT = renal replacement therapy; TSAT = transferrin saturation; URR = urea reduction ratio; v = variable; WBC = white blood count; y = annually; 2×/y = twice a year.
plete for conditions that are deemed relevant for daily dialysis practice and less so for other, less essential conditions. As a result, classification and validity of patients’ characteristics throughout the MONDO consortium encompass only broad categories but are fairly consistent across databases [14–16].

Missing Data
The primary databases almost completely capture the clinical course while the patient receives routine care. The providers allocate a considerable amount of effort to monitor each patient in their care with clearly defined, regularly repeated parameters in order to detect deterioration of the patient’s clinical condition or to monitor the quality of care delivered. Randomly missing data are therefore infrequent because patients are seen at short intervals and measurements can be repeated if they were accidentally forgotten or the patient missed a scheduled visit. However, these databases typically cannot capture periods spent outside the dialysis unit, such as hospitalizations, which are periods often associated with an adverse outcome.

Statistical Approach
MONDO provides a platform for a variety of observational studies, and the abundance and density of data call for innovative statistical approaches. Statistical analysis plans are developed collaboratively for each scientific project under the leadership of the Department of Statistics and Applied Probability, University of California Santa Barbara, Calif., USA. Which methods to use depends on the specific scientific questions, but, in general, they are required to deal with repeated measurements over time [17, 18]. For example, mixed-effect models may be used to account for heterogeneity among patients and correlation within a patient [19]. Heterogeneity and correlation at the facility level is also considered when relevant. Nonparametric regression methods may be required to deal with complex longitudinal patterns and to estimate trajectories [20]. When investigating longitudinal patterns and/or mortality with longitudinal measurements as time-varying covariates, a joint modeling approach is applied to reduce potential bias [21]. Since there is a large number of longitudinal measurements and none of them can fully predict the outcome, new statistical methods need to be developed to reduce dimensionality of high-dimensional longitudinal data that are also subject to informative censoring due to death. The large volume of data also poses computational challenges.

For example, in a recent publication, the characteristic dynamics before death for interdialytic weight gain, systolic blood pressure, serum albumin and C-reactive protein were investigated [2]. To estimate the mean functions and the trajectories of these variables, quintic spline models were fitted to partial conditional means. So far, the mean function for each variable was studied separately. Combining all variables with a goal towards building an alarm system that alerts the clinician to incipient deterioration of their patients’ condition remains an important research topic. Therefore, for the ultimate goal of improving outcome in dialysis patients, the MONDO initiative provides big challenges and ample opportunities for statistical method development.

Conclusion
By combining a number of primary electronic databases, the MONDO consortium has created a very large and diverse repository of patient level dialysis data with global outreach. Its main goals are to explore standard and novel monitoring parameters related to patient outcome, to develop new tools for risk prediction and to describe the clinical course of dialysis patients, including rare events and subcohorts with uncommon clinical presentations.

MONDO capitalizes on a worldwide practice to electronically monitor dialysis care prospectively and with parameter definitions that are very similar between providers [1, 8–10, 13, 22]. This results in the use of consistent methods of data collection between databases for a large majority of parameters [23, 24]. The methods for enrollment of study subjects are very similar in the primary databases because of a shared intention to represent patient care fully in the dialysis clinics. Inclusion is over 90% in all databases, resulting in a very small risk of selection bias associated with enrollment. However, it is possible that patients unable to give informed consent (e.g. due to advanced dementia, language problems or advanced illness), if not waived, are underrepresented in some of the cohorts. The scope of MONDO is broad and includes populations from areas which are not often included in international studies. This increases generalizability of results and the ability to detect rare events. However, it should be noted that primary databases are often not representative of their source population. Therefore, prevalence and incidence results in the MONDO cohort should be interpreted with caution. Another shortcoming of MONDO is that the relative sizes of the primary data-
bases result in an overrepresentation of patients from Europe and North America. Analyses are therefore usually first done by region in order to internally validate and confirm findings. The MONDO dataset lends itself to a variety of observational study approaches [20]. Key aspects are the abundance of longitudinal patient level data over long periods of time and a detailed description of the characteristics of each primary database. Since data collection is embedded in the routine care of the patients, recording of events occurring in the ambulatory setting is nearly complete. However, important gaps in follow-up occur when patients are hospitalized or treated outside the organization. Data definitions between databases are very similar and there is frequent and, most importantly, uniform follow-up, minimizing selection bias associated with diagnostic testing. New parameters that enter routine care are quickly available in the MONDO database. In summary, MONDO represents a dataset which holds great promise for a much more detailed understanding of the clinical course of patients on dialysis. It may serve as the foundation for the development of intelligent alert systems that can improve patient care.

Acknowledgments

We are indebted to the dedicated staff of the dialysis centers in the different countries for their efforts in collecting accurate data. A special thank you goes to all who helped converting primary data into the MONDO format: Paola Carioni, Qingqing Xiao, Peter Kaufmann and Eric Liu. We thank all MONDO consortium members for their immense contributions to this initiative.

Disclosure Statement

Drs. Grassmann, Kotanko, Levin, Marcelli, Toffelmire and Scatizzi hold stock in FMC.

References