Journal of Visualized Experiments

TBase – an integrated electronic health record and research database for kidney transplant recipients

--Manuscript Draft--

Article Type:	Methods Article - JoVE Produced Video
Manuscript Number:	JoVE61971R2
Full Title:	TBase – an integrated electronic health record and research database for kidney transplant recipients
Corresponding Author:	Danilo Schmidt, Dipl.-Inf. Charité - Universitätsmedizin Berlin Berlin, Berlin GERMANY
Corresponding Author's Institution:	Charité - Universitätsmedizin Berlin
Corresponding Author E-Mail:	danilo.schmidt@charite.de
Order of Authors:	Danilo Schmidt, Dipl.-Inf.
	Bilgin Osmanodja
	Matthias Pfefferkorn
	Verena Graf
	Dirk Raschke
	Wiebke Duettmann
	Marcel G. Naik
	Carolin J. Gethmann
	Manuel Mayrdorfer
	Fabian Halleck
	Lutz Liefeldt
	Petra Glander
	Oliver Staeck
	Michael Mallach
	Martin Peuker
	Klemens Budde
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US$2,400)
Please specify the section of the submitted manuscript.	Medicine
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Berlin, Berlin, Germany
Please confirm that you have read and agree to the terms and conditions of the author license agreement that applies below:	I agree to the Author License Agreement

Please provide any comments to the journal here.	
Please indicate whether this article will be Standard Access or Open Access.	Standard Access ($1400)

TITLE:

TBase – an integrated electronic health record and research database for kidney transplant recipients

AUTHORS AND AFFILIATIONS:

Danilo Schmidt[1], Bilgin Osmanodja[2], Matthias Pfefferkorn[2], Verena Graf[2], Dirk Raschke[1], Wiebke Duettmann[2], Marcel G. Naik[2], Carolin J. Gethmann[2], Manuel Mayrdorfer[2], Fabian Halleck[2], Lutz Liefeldt[2], Petra Glander[2], Oliver Staeck[2], Michael Mallach[1], Martin Peuker[1], Klemens Budde[2]

1. Division IT, Charité - Universitätsmedizin Berlin, Berlin, Germany
2. Department of Nephrology and Medical Intensive Care, Charité - Universitätsmedizin Berlin, Berlin, Germany

Email addresses of co-authors:

Bilgin Osmanodja	bilgin.osmanodja@charite.de
Matthias Pfefferkorn	matthias.pfefferkorn@charite.de
Verena Graf	verena.graf@charite.de
Dirk Raschke	dirk.raschke@charite.de
Wiebke Duettmann	wiebke.duettmann@charite.de
Marcel G. Naik	marcel.naik@charite.de
Carolin J. Gethmann	carolin.gethmann@charite.de
Manuel Mayrdorfer	manuel.mayrdorfer@charite.de
Fabian Halleck	fabian.halleck@charite.de
Lutz Liefeldt	lutz.liefeldt@charite.de
Petra Glander	petra.glander@charite.de
Oliver Staeck	oliver.staeck@charite.de
Michael Mallach	michael.mallach@charite.de
Martin Peuker	martin.peuker@charite.de

Corresponding authors:

Danilo Schmidt	danilo.schmidt@charite.de
Klemens Budde	klemens.budde@charite.de

KEYWORDS:

Electronic Health Records, Medical Informatics Applications, Health Information Management, User-Computer Interface, Data Collection, Databases as Topic, Software Design, Kidney Transplantation

SUMMARY:

TBase combines an electronic health record with an innovative research database for kidney transplant recipients. TBase is built upon an in-memory database platform, connected to different hospital systems, and used for regular outpatient care. It automatically integrates all relevant clinical data including transplantation-specific data creating a unique research database.

ABSTRACT:

TBase is an electronic health record (EHR) for kidney transplant recipients (KTR) combining automated data entry of key clinical data (e.g., laboratory values, medical reports, radiology and pathology data) via standardized interfaces with manual data entry during routine treatment (e.g., clinical notes, medication list, and transplantation data). By this means, a comprehensive database for KTR is created with benefits for routine clinical care and research. It enables both easy everyday clinical use and quick access for research questions with highest data quality. This is achieved by the concept of data validation in clinical routine in which clinical users and patients have to rely on correct data for treatment and medication plans and thereby validate and correct the clinical data in their daily practice. This EHR is tailored for the needs of transplant outpatient care and proved its clinical utility for more than 20 years at Charité - Universitätsmedizin Berlin. It facilitates efficient routine work with well-structured, comprehensive long-term data and allows their easy use for clinical research. To this point, its functionality covers automated transmission of routine data via standardized interfaces from different hospital information systems, availability of transplant-specific data, a medication list with an integrated check for drug-drug interactions, and semi-automated generation of medical reports among others. Key elements of the latest reengineering are a robust privacy-by-design concept, modularity, and hence portability into other clinical contexts as well as usability and platform independence enabled by HTML5 (Hypertext Markup Language) based responsive web design. This allows fast and easy scalability into other disease areas and other university hospitals. The comprehensive long-term datasets are the basis for the investigation of Machine Learning algorithms, and the modular structure allows to rapidly implement these into clinical care. Patient reported data and telemedicine services are integrated into TBase in order to meet future needs of the patients. These novel features aim to improve clinical care as well as to create new research options and therapeutic interventions.

INTRODUCTION:

Motivation for an integrated electronic health record and research database

Clinical research is based on the availability of high-quality data, regardless of whether classical statistical methods or Machine Learning (ML) techniques are used for analysis[1,2]. In addition to routine data (e.g., demographic, laboratory, and medication data), domain-specific data (e.g., transplantation-relevant data) are required with high granularity[3,4]. However, routine care at many university hospitals is performed with hospital information systems (HIS) that neither allow for systematic collection of research-specific data nor for easy data extraction of routine data[5-7]. As a result, clinical researchers create specific research databases, which have a variety of problems including complex process of setting up a database, manual data entry, data protection issues, and long-term maintenance (**Table 1**). Limited amount of data, missing data, and inconsistencies are a major problem for clinical research in general and impede the use of ML technologies[8-13]. These standalone research databases are usually focused on certain disease or patient aspects, not connected to other databases, and often discontinued after a certain period, resulting in inaccessible “data silos”. Ultimately, high-quality, long-term data on various disease aspects are sparse. In the era of digital medicine there is an increasing need for a comprehensive electronic health record (EHR)[7,14,15], which enables easy documentation of domain-specific data and automated collection of routine data from the systems of inpatient and outpatient care.

These general considerations apply to transplantation medicine as well[16]. Hence, a complete documentation of the patient's medical history including all inpatient and outpatient treatments, clinical routine data as well as transplantation-specific data is necessary for successful follow-up care[17,18]. Since ordinary HIS are static and focused on inpatient treatment, they cannot integrate transplantation-specific data, such as donor data, cold ischemia times, and human leukocyte antigens (HLA) data. However, these data are a basic prerequisite for transplantation research[19-22] as well as from long-term clinical care. While the initial hospital stay usually is only 1-2 weeks and processes as well as early outcomes after kidney transplantation are comparable between many transplant centers, lifelong post-transplant care is complicated and lacks a common structured approach. This motivates an integrated EHR and research database to capture the lifelong post-transplant patient journey.[23]

In order to integrate these functionalities for routine care and research of KTR, an EHR named "TBase" was developed with the idea that the routine use for post-transplant care will create a unique research database with highest data quality (**Table 2**).

Design and Architecture

TBase is based on a typical client-server architecture. For development, the components and tools of SAP High Performance Analytic Appliance extended application advanced (SAP HANA XSA) were used. Based on the latest Hypertext Markup Language 5 (HTML5) web-technologies the EHR has been developed and tested for the Google Chrome Engine. This web engine is used by Chrome and the Microsoft Edge Browser and allows to use the EHR in the most frequently used web browsers[24] without the need for local installation. The applied technology enables a responsive web design and allows the web-based EHR to be used on all devices (PC, tablet, smartphone). The innovative high-performance development platform is composed of various components (Web IDE, UI5 and HANA DB) and has enabled us to rapidly implement the EHR project TBase with state-of-the-art software tools (**Figure 1**).

For the representation of patient data, a simple table structure was implemented for an intuitive and self-explanatory design of the EHR. For example, the patient table with the PatientID as the primary key is at the center of the table structure. Almost all tables (except individual sub-tables) are connected to this central table through PatientID (**Figure 2**).

Figure 3 shows part of TBase's table structure and the data types used in greater detail. The end user can access the data fields via graphical user interface (GUI), for which an example is shown in **Figure 4**.

This EHR contains all current patient data and is used for routine outpatient care. Important routine clinical data (e.g., laboratory data, medical results, radiology, microbiology, virology and pathology data, hospital data, etc.) are directly imported into TBase via standardized interfaces (e.g., on the basis of Health Level Seven (HL7) – a standard for digital communication in the healthcare sector[25]). Transplant-specific data such as cold ischemia times, donor data, HLA data as well as follow-up notes, vital signs, medical reports and the medication list are entered by the

users via GUI into the EHR. Before data are transferred to the database, an automated plausibility check is performed for prompt detection of erroneous data entry providing the option to correct immediately. In addition, data validation takes part during clinical routine in which clinical users routinely write reports and letters to patients and physicians. These letters must provide correct data (e.g., on medication, lab values and clinical remarks) for further treatment and medication plans. As a consequence physicians and patients constantly validate and correct the clinical data in their daily practice, a process resulting in high data quality. If data are entered via application programming interfaces (API) or other interfaces, plausibility checks are performed in the backend similarly to the plausibility checks in the frontend.

Frontend (GUI)

To implement the frontend, the UI5 Framework is used. This framework provides an extensive library for frontend elements as well as a variety of additional features such as multilingualism and graphical libraries for data visualization. Currently, TBase frontend elements are displayed either in English or German depending on the language setting of the browser.

A master-detail interface is used for the frontend to ensure a simple, intuitive page structure. The upper part of the viewing page consists of individual tabs for the detail pages (basic data, medical data, transplantation data, etc.). This master part remains unchanged regardless of which detail page is shown below (**Figure 4**). The detail view of each page enables an easy overview on the page topic.

For data manipulation, the EHR has different levels of user rights ("read", "write", "delete", and "administrator"). There is an "edit" level in addition to the "view" level, which can only be activated by users with higher rights than "read". If the user has the right to write, all input fields for data entry are activated and can be filled with data. Users with “delete” rights can delete data via a corresponding button, but only after confirmation through a pop-up window.

Database structure and interfaces

The development of TBase is performed in the development database. Extensive and detailed testing of all software changes such as new functionalities is carried out in the quality assurance database. Software updates that pass the quality control checks are transferred to the live system. For research purposes the live system is copied into the replication database, which can be queried via standard Open Database Connectivity (ODBC) interfaces (e.g., via open-source software R Studio). As there is no direct connection between replication and live system the data in the live system are protected from corruption, loss or manipulation of data. This modular structure and the clear separation of the four databases (development, quality assurance, live system and replication database), which are tailored to the specific needs of developers, researchers, and clinicians facilitates maintenance and data protection of sensitive patient data.

The EHR is fully integrated into the Data Infrastructure of Charité and relies on different interfaces for data import from various data sources. The interface to the HIS imports all relevant data such as administrative data, examinations, medications, laboratory findings and discharge letters. This interface connects both systems via a staging area. Here, all new data (data delta) are transferred

from the HIS to TBase in real time. Patients are identified via a patient number or case number and the corresponding data from the HIS is imported (if not already available in TBase).

For outpatients, our laboratory partner provides the laboratory results via HL7 messages. These are deployed to a shared area in the laboratory system and picked up via an HL7 interface and imported into the EHR. For bi-directional communication and data exchange with KTR (via smartphone apps) and home nephrologists, a HL7 Fast Healthcare Interoperability Resource (HL7 FHIR) interface was implemented[26]. This interface grants interoperability and flexibility for a safe data exchange with other data sources (e.g., Eurotransplant, patient apps) in the future.

User Management and Data Protection

TBase is based on user management at the application level. Thus, the user can only access the frontend of the application, but not the database itself. As described above, a four-stage authorization concept was chosen, reserving user management for those with administrative rights. Administrators use an "Identity Management Console" application to add new users from the Charité user pool for the TBase application and to maintain their user rights (**Figure 5**). Most users can access all patients in the database. However, it is possible to restrict access for specific users such as study monitors to a group of patients.

Using the commercial in-memory database platform a secure database technology that protects data with strategies such as application-level authorization, single sign-on (SSO), MIT-Kerberos protocol and Security Assertion Mark-up Language (SAML) is used. The platform secures communication, data storage, and application services using the latest encryption and testing techniques. All developments on the database are controlled by authorizations. This ensures the security of data by design at a high level. In addition, all data are kept behind the certified Charité firewall. In compliance with the latest European Union General Data Protection Regulation (EU GDPR) a robust data protection concept was implemented, including data flow diagrams, data protection risk assessment (DSFA) and authorization concept. All documents are laid down in a procedure directory of the Charité Data Protection Office.

PROTOCOL:

The protocol demonstrates the use of the electronic health record TBase, how to add data into the database, and how to extract them for research purposes. All steps are in accordance with the guidelines of the human research ethics committee of Charité – Universitätsmedizin Berlin.

1. Register a new patient and add basic patient data into TBase

1.1 Upon registration, transfer the patient's basic data (name, birth date, and health insurance data) from the patient's health insurance card to the hospital information system. During this process, a new unique case number is created. If the patient has never been treated at Charité – Universitätsmedizin Berlin, a new unique patient number is created as well, that clearly identifies this particular patient in the hospital system.

1.2 During this registration process, obtain written informed consent from the patient for TBase data processing by Charité – Universitätsmedizin Berlin and the outpatient clinic of Charité (Ambulantes Gesundheitszentrum der Charité) according to the EU GDPR.

1.3 Have an employee with appropriate permission add this new patient to TBase. First, sign into TBase via the GUI. For that, enter "https://nephro.tbase.charite.de" into a Chrome-Engine based Web Browser in Charité Intranet. Next, enter the username and password assigned by the TBase administrator. Click **Log On**.

1.4 Next, click on the **Add new patient** button on bottom of the **Patient** overview frame on the left. Then, an input screen appears.

1.5 Enter the patient's name, birth date, Charité hospital patient number (see above, or alternatively a Charité hospital case number), and the information about patient data processing consent (if it is granted, not granted or revoked by the patient). Click on the **Save** button on the bottom right when data entry is completed.

NOTE: Now, a new patient has been added to TBase and automatically all available patient data are now transferred from HIS into the TBase EHR system.

2. Viewing and adding data to a patient record in TBase in the Sections: Master Data, Medical Data, Doctors, Diagnosis, Procedures, Transplantation Data, Hospital

2.1 Log in to EHR as described in step 1.3.

2.2 Search for the desired patient via the search field on the top left via name or birth date. Click on the search button right of the search field or hit **Enter**. From the results in the **Patient** overview frame on the left, choose the right patient and click on the name. A new screen appears, showing the selected patient's **Master Data**.

2.3 After searching for a patient, the patient's **Master Data** viewing page appears by default. To navigate there from another page, click on the **Master Data** tab on the top left.

2.3.1 To change **Master Data**, click on the **Change** button on the bottom right. A new input screen appears.

2.3.2. Now, change data such as patient's phone number, address, add or correct an identification code by typing the new information into the designated input fields. After data entry is completed, submit the changes by clicking on the **Save** button on the bottom right. After being redirected to the **Master Data** viewing page, changes can be seen and verified.

2.4. To view and change **Medical Data**, click on the **Medical Data** tab on the top left. The **Medical Data** overview appears and shows the existing medical data. They are structured as follows: patient's height, blood type, first dialysis date, primary disease, HLA, genetics data, dialysis data,

data on existing HLA-antibodies, transfusion data, risk factors, allergies, structured anamnesis data, death.

2.4.1 To change some of the medical data, click on the **Change** button on the bottom right. A new input screen appears.

2.4.2 For example, add a primary disease to the patient's medical data by clicking on **Primary disease** to expand or collapse the data entry form. On the right, the **primary disease** input field can be used to select one disease out of the preexisting suggestions (e.g., from Eurotransplant primary disease table) or to enter a new disease. Additionally, information about the diagnosis date, the certainty of the disease (biopsy-proven or not) and a comment can be entered. After data entry, submit by clicking on the **Submit Values** button.

2.4.3. After all changes have been entered and submitted, save the changes by clicking on the **Save** button on the bottom right. After being redirected to the **Medical Data** viewing page, all changes can be seen and whether they have been saved correctly.

2.5. To view information about the treating physicians, click on the **Doctors** tab. The **Doctors** viewing page opens up and shows the existing data about treating physicians. They are structured as follows: physician's name and address, specialization, type (consultant, general practitioner, resident), working facility (dialysis ward, outpatient clinic, etc.), phone number.

2.5.1 To add a new physician, click on the **New** button on the bottom right. A new input screen opens up. Alternatively, information about existing physicians can be modified by clicking on the physician's name first, and afterward clicking on the **Change** button on the bottom right.

2.5.2 For example, a new physician can be added to the patient's EHR. Search the list of previously added physicians by entering a name into the search field and clicking on the right entry from the different suggestions. Alternatively, if the desired physician is not in the list, enter the data into the input field below after selecting **Add New Physician** first.

2.5.3. After all changes have been entered, save the changes by clicking on the **Save** button on the bottom right. After being redirected to the **Doctors** viewing page, all changes are visible and the user can verify that the changes have been applied correctly.

2.6. To view and change diagnoses, click on the **Diagnosis** tab on the top left.

NOTE: Most of the diagnoses, procedures and investigations are automatically imported via predefined interfaces from the HIS about inpatient treatment data.

2.6.1 Enter diagnoses made in the outpatient clinic by clicking the **New** button on the bottom right.

2.6.2 A new diagnosis can be entered, based on International Classification of Diseases 10: Revision

(ICD-10). Enter the ICD-10 code or the diagnosis name into the search field in the center of the screen and select the right one from a suggestion list by clicking on it. Next, define the start and end date if applicable and the context, where the diagnosis was made (inpatient or outpatient) by typing these data into the designated input fields.

2.6.3 Submit the data by clicking the **Save** button on the bottom right. After being redirected to the **Diagnosis** viewing page, the changes become visible and the user can see, whether data entry was correct.

2.7. To view and change procedures, click on the **Procedures** tab on the top.

2.7.1 Enter additional procedures performed in the outpatient clinic by clicking the **New** button on the bottom right.

2.7.2 A new procedure can be entered, based on OPS-Code (German version of International Classification of Procedures in Medicine (ICPM) codes). Enter the OPS-code or the procedure's name into the search field in the center of the screen and select the right one from a suggestion list by clicking on it. Next, define the localization (left, right, none) and the context, where the procedure was performed (inpatient or outpatient) by typing these data into the designated input fields.

2.7.3 Submit the data by clicking the **Save** button on the bottom right. After being redirected to the **Procedures** viewing page, verify that the changes have been applied correctly.

2.8 To view and change data on investigations, click on the **Investigations** tab on the top.

NOTE: Since most of the reports in the HIS are provided as text-files, most of the corresponding results in the EHR are text-based as well. In contrast, pathological reports from kidney transplant biopsies are classified according to Banff Classification 2017[27,28] and the resulting discrete classification data are saved into a corresponding table in the EHR.

2.8.1 To look at the findings of a specific investigation click on the right one in the list below or use the search field above to select it from the suggestion list.

2.8.2 Enter additional investigations performed in the outpatient clinic by clicking the **New** button on the bottom right.

2.8.3 Enter a new investigation by typing date, type (ultrasound, holter-monitoring, etc.), involved organ and the findings into the designated input fields.

2.8.4 Submit the data by clicking the **Save** button on the bottom right. After being redirected to the **Investigations** viewing page, the changes can be seen and verified by the user.

2.9 To view and change data on hospitalizations, click on the **Hospital** tab on the top.

NOTE: Regularly, KTR that have been transplanted at Charité are hospitalized at the transplant center for subsequent complications. Generated data is firstly stored in the HIS and relevant data (e.g., data about admission or discharge, medical reports) are imported into EHR via HIS interface. External hospitalization have to be entered into EHR manually.

2.9.1 The data on hospitalization is structured as follows: admission, discharge, medical report if available, hospital, ward, and reason for hospitalization. To read the medical report, click on the right one in the list or use the search field above to select it from the suggestion list.

2.9.2 Enter an additional hospitalization (e.g., external hospitalization) by clicking the **New** button on the bottom right.

2.9.3 Enter a new hospitalization by typing above-mentioned data into the designated input fields.

2.9.4 Submit the data by clicking the **Save** button on the bottom right. After being redirected to the **Hospital** viewing page, where the changes become visible and can be verified.

2.10. To view and change transplantation data, click on the **Transplantation** tab on the top right. The **Transplantation** viewing page appears and shows the existing transplantation data. On the top, navigate between different transplantations by clicking on the corresponding button, if more than one transplantation has been performed.

2.10.1 To view or change information about the donor, click on the **View Donor** button below the corresponding transplantation date. To enter or change information about the donor, click on the **Change** button at the bottom right and enter data into the designated input fields and save the changes by clicking on the **Save** button on the bottom right thereafter.

2.10.2 To add a new transplantation to the patient's EHR, click on the **New** button on the bottom right on the **Transplantation** viewing page. Enter transplantation specific data according to the input fields (including information about organ type, transplantation date, ischemia time, procedural complications among others). Save the data to the EHR by clicking on the **Save** button at the bottom right. The user is then redirected to the **Transplantation** viewing page to see whether the changes have been saved correctly.

2.10.3. To change information about an existing transplantation, click on the **Change** button on the bottom right on the **Transplantation** viewing page, and a new input screen appears where the existing data for the selected transplantation are shown. Change these transplantation specific data according to the input fields (including information about organ type, transplantation date, ischemia time, procedural complications among others). Save the new entry data in the EHR by clicking on the **Save** button at the bottom right. After being redirected to the **Transplantation** viewing page, see the changes and check whether changes are entered correctly.

3. Viewing and selecting laboratory data

3.1. Log in to TBase and select the desired patient as described in 1.3 and 2.2.

3.2. To view the laboratory data, click on the **Laboratory** tab on the top, and a tabular overview of the latest laboratory results appears. On the top, all data of the last investigation are visible with a drop-down menu to search for previous lab data and a search field next to it, where one can search for specific laboratory values (e.g., creatinine).

NOTE: The laboratory values are displayed as follows: date of sample receipt, date of processing, name of the laboratory value, value, unit, reference range, a comment (H … high, L … low, N … normal), and the previous two historic lab values for comparison.

3.3 To change a date for view of a historic laboratory investigation, click on the drop-down menu on the top left and select the desired date by clicking on it. All corresponding lab values from this date is then displayed as described above.

3.4 To select a specific laboratory value such as creatinine and examine its course over time, type its name into the search field on the top and select the right one from the suggestion list. After clicking the **Show Labor** button, every result for the selected value of this patient is shown in the chart below.

3.4.1 Alternatively, simply click on the desired value in the initial tabular presentation of a single investigation. This again shows all previous and the current results for this specific laboratory value.

3.5 To plot the course of a laboratory value, click the plot symbol next to the desired value. This automatically creates a plot of all existing results for this value. If needed, specify the time range for the plot by selecting a start and end date in the input fields on the top right and add a second value to the plot by selecting it in the designated input field. Go back to the **Laboratory** viewing page by clicking on the **Back** button on the bottom right.

4. Viewing and changing medication data: creating a standardized medication list according to German regulations (“Bundeseinheitlicher Medikationsplan”)

4.1 Log in to TBase and select the desired patient as described in 1.3 and 2.2.

4.2 To view the medication data, click on the **Medication** tab on the top. A tabular overview about the patient’s current medication appears. The medication data are shown as follows: starting date, active substance, single dose (e.g., in mg), trading name, dosing scheme, daily dose, dosage form, notification, indication, kind of prescription (internal or external physician, or self-treatment by the patient).

4.3. To add a new medication, click on the **New** button at the bottom right. Enter the name of the substance (or alternatively the trade name), the dosing scheme, and the starting date, which is set automatically at the current date, but can be changed if the starting date was in the past. Additionally, indication and a remark can be added into the designated input fields. Add the medication to the list by clicking on the **Save** button on the bottom right.

4.4 To change an existing medication, click on the appropriate item in the medication list and on the **Change** button on the bottom right afterwards. Now, changes regarding dosage, application form can be typed into the designated input fields and the changes can be applied by clicking the **Save** button on the bottom right.

4.5 To discontinue a drug, click on the designated drug and click on the **Discontinue** button on the top.

4.6. To search for previous medication, enter the active substance into the search field on the top left and select the right one by clicking on it from the historic medication list. A chart with all previous medications appears, which is structured as stated in 4.2.

4.7. To create a standardized medication list for the patient according to German regulations, click on the button **Bundeseinheitlicher Medikationsplan** on the top right. A PDF-file is created and downloaded automatically for printout.

5. Viewing and adding entries to the medical course: generating a medical report semi-automatically

5.1 Log in to the EHR and select the desired patient as described in 1.3 and 2.2.

5.2 To view the medical course, click on the **Course** tab on the top. A tabular overview about the documentation from the patient's previous appointments is provided. The data are structured as follows: date of the appointment, date of the next appointment, blood pressure, heart rate, temperature, weight, body mass index, urine volume and three text fields divided into public assessment for the patient, internal assessment for use at Charité, and medical assessment for other physicians.

NOTE: Additionally, there is a summary field at the bottom, which is used to summarize important information about the patient's medical history and make it visible at first sight.

5.3 To add a new entry to the medical course, click on the **New** button on the bottom right. Enter the information assessed into the desired input fields (e.g., vital signs, treating physician, internal assessment or public assessment). Add the date of the next appointment into the designated input field on the top right. Submit the data by clicking the **Save** button on the bottom right. Users are then redirected to the **Course** viewing page.

5.4 To change an existing entry, click on the appropriate one and click the **Change** button on the

bottom right next. Now, enter additional data into the designated input fields or change existing data. Change or update information in the notification field by typing into it, and submit the changes by clicking the **Save** button on the bottom right. Users are then redirected to the **Course** viewing page.

5.5 To create an automated medical report, click on the **Medical Report** button on the bottom right. A new screen appears, with 18 different options (ranging from laboratory results to complete medical report).

5.5.1 For example, create a medical report with a few clicks: Click on **Outpatient Medical Report**. The patient name, the treating physician, the last date of the laboratory values and last date of medical course are automatically filled out, but can be changed if needed. After confirmation by clicking on **OK,** a properly formatted word (.doc-) document file is created and downloaded for printout containing the selected information.

6. Logging out

6.1. To actively log out of TBase, click the **Log out** button at the bottom right. Additionally, one is logged out automatically after 60 minutes of inactivity or if the browser is closed.

7. Using the collected data

7.1. To query the collected data, use replication server (**Figure 1**) as described in the Database structure and interfaces section. Any data processing programs that can connect to a database via Open Database Connectivity (ODBC), Java Database Connectivity (JDBC) can be used for the queries. Once the connection to the database has been established, use the open-source software R Studio.

7.2. To set up an ODBC database connection, for example, in the Windows operating system, open the ODBC tool and click **Add** for a new user data source name (DSN) under Control Panel and Security Management. There, enter the available connection data to the replication database. Enter the following data: "Driver Name", "ODBC Connection Name" (set by the user), "Hostname" and the SQL authentication details "User Name", "Password" and "Database Name".

7.3. In order to generate a very simple query (e.g., number of transplants divided by gender in the years 2000-2020) in the open-source software R Studio after the ODBC database connection has been set up, open **File**, **New File** in the application **R Studio** at the top left and click on **R Script**. The example script code (**Code 1**) is entered in the empty script window that opens.

7.4. Click on button **Source** on the top of the script window and the script is running and then generates the bar chart defined in the script with the data from the connected database (**Figure 6**).

REPRESENTATIVE RESULTS:

TBase was first released in 1999 at Charité Campus Mitte and is in use ever since. For more than 20 years the TBase-EHR prospectively collects data from all KTR. Starting in 2001, the other transplant programs at Charité used TBase for the routine care of KTR and wait-listed patients as well. Since 2007, this EHR is in use for routine care of living donors and all patients in the department of nephrology.

By providing the TBase software with its functionalities, which has been further developed in recent years into a modular web-based research database with modern software architecture, a total of 6,317 patients with 7,595 kidney transplantations were documented over a period of more than 20 years. In total the KTR have 220,877 diagnoses, 332,299 procedures, 1,033,941 laboratory reports, 24,478,441 Laboratory values, 539,922 medication episodes, 324,339 investigations, 6,489 donor data, and 54,350 discharge letters among others (**Table 3**).

In addition, data of 20,724 patients were collected including patients on the waiting list, living donors and patients with chronic kidney disease. These patients have a total of 232,783 diagnoses, 408,857 procedures, 546,661 laboratory reports, 13,399,048 Laboratory values, 114,657 medication episodes, 226,206 investigations and 70,278 discharge letters.

More than 50 scientific publications from the TBase database as original research in peer-reviewed journals were published over the last 10 years[1-4,6,9-13,17-19,21,29-36]. A thorough data protection assessment was performed and laid down at the data protection officer. TBase is supervised by a development team consisting of four fulltime computer scientists with the support of IT department and the department of Nephrology.

FIGURE AND TABLE LEGENDS:

Table 1: Problems of clinical databases and standalone research databases. This table lists the problems of hospital information systems (HIS) on the one hand and specific research databases with a variety of problems of the other hand.

Table 2: Benefits of TBase in clinical care and research. This table lists the benefits of an integrated EHR and research database for clinical uses and researches.

Table 3: Number of patient records and patient data. For more than 20 years TBase prospectively collects data from all KTR. Starting in 2001, the other transplant programs at Charité used TBase for the routine care of KTR and wait-listed patients as well.

Figure 1 - TBase architecture. The EHR TBase is based on four different databases with the live system at core enabling different groups of agents to work in parallel. Clinical users enter data via the graphical user interface (GUI) and maintain the high data quality. Most information is automatically imported via interfaces from hospital information systems, laboratory partners, and drug-drug interactions checkers. Developers can implement new functionalities, which are tested on the quality database before integrating them into the live system. For research purposes the live database is replicated regularly, so that no interference with the live system is

necessary when database queries are performed by clinical researchers.

Figure 2: TBase data tables. For the representation of patient data, a simple table structure was implemented for an intuitive and self-explanatory design. For example, the patient table with the PatientID as the primary key is at the centre of the table structure. Almost all tables (except individual sub-tables) are connected to this central table through PatientID.

Figure 3: TBase data types. To exemplify how different key clinical data are represented in the database underlying TBase, parts of the tables Medication, Patient, and Transplantation with the corresponding data types are shown. The data fields are represented in the TBase GUI. For instance the **Medication** table is typified in the **Medication** mask.

Figure 4: TBase screenshot. on top the **Master** view displays patient name and different tabs for different **Detail** masks, here the **Medication** mask. On the left drug-drug interactions are displayed by color codes. In each line, start date, drug and trading name, dosing schedule, potential notes and the indication for prescription are shown. Additional buttons in the mask and below indicate different functionalities such as historic search, start and stop of medication, as well as print of medication plan for the patient.

Figure 5: Authorization concept. The EHR is based on user management at the application level. Thus, users can only access the frontend of the application, but not the database itself. A four-stage authorization concept was chosen reserving user management for those with administrative rights. Administrators use an "Identity Management Console" application to add new users from the Charité user pool for the TBase application and to maintain their user rights.

Figure 6: Result of a very simple database query. The database can be connected via a database interface (e.g., ODBC). Data processing software (e.g., the open-source software R Studio) can be easily sent queries to the database and graphical results can be generated.

Figure 7: TBase on a professional modern in-memory database system. Using the development system facilitates the simultaneous development of TBase applications in different development containers for other user groups with their unique domain requirements. The specific containers for different user groups run in parallel and are brought into the live system via the TBase Master Template.

DISCUSSION:

TBase combines a web-based EHR for specialized outpatient care of KTR with a research database, creating a comprehensive long-term database for patients with kidney disease[6,11,15,37]. Regarding organizational structure, this is enabled by implementing a modern software design process as an institutional agent and including over 20 years of experience as developers, clinical users and researchers to develop the current version. Additionally, it is constantly improved and updated according to the needs of the clinical users, since it is the main outpatient documentation software for nephrological patients at Charité - Universitätsmedizin Berlin. For the clinical user, this EHR offers automated integration of all relevant data from inpatient and

outpatient systems and allows for easy and intuitive documentation of transplant-specific data and long term patient pathways. Regarding medication data, a commercially available database is connected to the EHR, providing information about every pharmaceutical product approved in Germany, which is updated every two weeks. Prescribed medication is checked for drug-drug interactions by cooperating with a commercial system from the Department of Clinical Pharmacology & Pharmacoepidemiology at the University Hospital Heidelberg through an API. With each medication change in TBase, the respective medication data are sent to the system in a pseudonymized form. The software identifies potential drug-drug interactions of the active agents. The results are sent back to the EHR, where they are stored and displayed in the medication mask in real time. A color code marks the severity of potential interactions and a pop-up window provides detailed information on the type of interactions (**Figure 4**). The patient can access the same medication data as a standardized medication list as part of a semi-automated medical report or via live updates on the smartphone[16].

Regarding its database architecture, the EHR is based on four different databases with the live system at core, hence enabling different groups of agents to work in parallel. It is practice-approved that development, testing, and clinical work do not hinder each other. For research purposes the live system is replicated weekly but not automatically, which to improve is under development. Most importantly, the live system is separated from developers and researchers providing highest level of data protection for the sensible patient data by design. In addition, by this means data manipulation and corruption is limited and data validity in maintained as much as possible.

Technologically, the EHR is based on a modern in-memory database technology consisting of various components and guaranteeing secure communication and data storage. In the backend, patient data are protected by a granular authorization concept. For that purpose, the development platform is separated from the live system and different authorization levels are deployed. Additionally, all data are saved on Charité servers behind the certified firewall, further increasing data protection.

Regarding the frontend, the EHR is based on HTML5 technology, which enables responsive web design and allows to use the web-based EHR on all end devices at any place in the hospital. The development environment supports numerous standard programming languages such as Java, Javascript, PHP, and Python, thus enabling easy and successful recruitment of developers. Furthermore, numerous pre-designed visualization options are available, generating graphics for the course of laboratory parameters as described in the protocol. In the next developmental steps real-time monitoring of patient data (e.g., renal function, vital signs) and important outcomes such as patient and graft survival will be integrated. Automated display of waitlisted patients and their current status will improve process management regarding organ allocation from Eurotransplant. For this purpose, interfaces to external partners (physicians, dialysis wards, Eurotransplant) are necessary, as discussed next.

One primary goal when creating TBase was to bring all outpatient and inpatient routine data into one database. This is enabled by APIs and self-developed interfaces, which, to this point, import

different kinds of laboratory results and other investigations from inpatient and outpatient treatment, while new interfaces to Eurotransplant and other external partners are under development. Regarding laboratory data for outpatients, our laboratory partner provides the laboratory results via HL7 messages. These are deployed to a shared area in the laboratory system and picked up via an HL7 interface and imported into the TBase system. Since it is an external provider and does not have access to the hospital information system, the Charité hospital patient number or case numbers cannot be used for patient identification. So far, the only other possibility for identification via laboratory order numbers could not be used either, because the electronic laboratory orders are generated in a separate system and the EHR does not have automatic access to these. For this reason, incoming data can only be automatically identified by name, first name and date of birth. Since spelling in the different systems may differ, it occasionally happens that data cannot be automatically matched in TBase. These observations are cached in a queue until they are assigned by a user manually to the correct patient in the EHR. The assigned patient identifiers are stored in the system and will be used for automatic identification in the future.

An important concern when designing clinical databases is changing data quality over time. Since the effects depend on the type of data considered, we deliberate this for different data types independently, based on those mentioned in **Table 3**. Most data in TBase are routine laboratory data. While laboratory methods may have changed, the parameters themselves remained the same throughout time, such as creatinine, urea, electrolytes, liver enzyme assays, coagulation parameters, and blood count. The parameters and the corresponding normal ranges are comparable over time, even if the latter differ slightly between different assays. At our laboratory, most laboratory measurement do not use SI units, but metric units. While this requires to implement conversion factors when incorporating external data, it ensures comparability of internal laboratory data over the last 20 years. The second largest group are medication data. Despite the fact that standard immunosuppressive regimes have changed, data quality about medication did not change over time. Procedures are encoded as OPS-codes, which are the German adaption of the International Classification of Procedures in Medicine (ICPM) and are universally accepted and used ensuring continuous data quality over time. The same holds true for diagnosis data, which are encoded according to ICD-10 that is obligatory for outpatient medicine in Germany since 2000. Another big group of data are those about medical investigations. In kidney transplantation, especially data about kidney transplant biopsies are important. For pathology data of kidney transplant biopsies, currently the Banff 2017 classification is used, but classification for kidney transplant biopsies have changed multiple times over the last 20 years. We therefore decided to reclassify all kidney transplant biopsies made at Charité Berlin according to Banff 2017 to remarkably increase data quality for research purposes[27,28]. Other medical data such as imaging reports or those from examinations such as colonoscopy or bronchoscopy are recorded as strings. To restructure those data according to the latest medical standards and to make it available for research purposes, we currently work on categorizing such data retrospectively. The same applies to medical records. Their formatting and structure vary among different departments and over time, so we currently decided to record them as strings. This enables us to use the incorporated information in the clinic and for research purposes. The last group of data, are patient data such as age, gender, treating physicians, but

also general medical data such as allergies, duration and form of dialysis. These did not change remarkably over the last two decades. There is, indeed an increase in data quality about underlying kidney diseases with greatest progress made in the detection of genetic diseases. Rarely, patients who are registered for a second transplantation are tested for genetic diseases and their diagnosis is corrected subsequently, but for most patients such correction cannot be achieved. Similarly, progress has been made regarding transplantation data. Especially detection of donor-specific anti-HLA antibodies is a highly dynamic field of research with source data from HLA-laboratories filling their own databases. Hence, it is only recorded, if donor-specific HLA-antibodies are present. Therefore, changes in detection method are not considered. Overall, for most data types, we already ensure homogenous data quality over time. In the future, we will implement methods like Natural Language Processing to extract relevant information from text based data like medical reports or examination reports.

Outlook – TBase User Group, Artificial Intelligence, and Telemedicine

Since TBase is designed in a modular way, scalability into other transplant centers and to different outpatient clinics in other hospitals is possible and under current development (**Figure 1**). Especially other transplant centers and specialized outpatient clinics of Charité are now implementing the EHR for their individual clinical and research purposes. The basic structure is maintained and extensions tailored to the individual needs are programmed for endocrinology, rheumatology, neurology, cardiology and gastroenterology. One main advantage for the different departments is TBase's integration into the IT landscape of Charité with a state-of-the-art data protection concept. The different TBase containers were implemented in a short time thanks to the modular structure and the existing interfaces (**Figure 7**). Recently, a database for follow-up and clinical research of COVID-19 patients based on TBase was established as well.

Sharing and developing the integrated database with other departments and hospitals helps to build a more sustainable research platform that is open to other transplant centers. The goal of such an open source developmental process is to create synergies via a TBase User Group. This raises new problems of varying data protection measures at different institutions, but offers the possibility to enhance further development by including other institutional agents with additional expertise as well. The user group can join forces and implement new modules (e.g., for clinical trials) or interfaces (e.g., for Eurotransplant, quality assurance, transplant registry). Another important perspective for TBase is the certification as a medical device, which can be better achieved if several partners work together. The aim of the TBase User Group is to create a flexible platform that each department can use on its own, but which is also available for larger common clinical research projects.

Another important aspect worth considering is that integrating algorithms, especially Machine Learning technologies, into medical workflows is most promising, when performed on high quality health data[16,38]. TBase seems well-suited for the development as well as the implementation of novel AI methods in kidney transplantation by a variety of captured variables, the availability of structured and unstructured data as well as high-quality long-term data. Future AI-based clinical decision support systems can be integrated seamlessly into the Graphical User Interface of the EHR. Before such systems can be deployed, clinical, technical, ethical and legal

challenges have to be overcome. First of all, the clinical utility of such AI-based algorithms has to be proven. Therefore, we currently conduct a study, in which we test the accuracy of an AI-based algorithm for predicting rejection, transplant graft failure or infection in KTR. The algorithm was developed by the German Research Institute for Artificial Intelligence (DFKI) based on TBase data using Gradient Boosted Regression Trees. The study will compare the ability of experienced physicians to predict different endpoints with and without information provided by the AI-model. In a next step, we will apply text mining and methods from the field of Natural Language Processing. Hereby, we aim to create structured data from medical reports and examination reports to make this huge body of valuable data available to research and the clinician. This will later enhance model accuracy of AI-based prediction models, which are currently not incorporating these unstructured data. Beyond those technical and medical consideration, ethical and legal challenges have to be met. Therefore, we additionally examine human-machine interaction in the abovementioned experiment to understand what opportunities, but also what risks and disadvantages arise when implementing such systems from the physicians' perspective. Therefore, we conduct structured interviews and examine how the use of such decision support systems changed the physicians' approach to medical decision making.

A third forward looking modular expansion of TBase is to integrate telemedicine solutions. In contrast to many telemedicine approaches, which cannot be transferred into the existing digital infrastructure, complete integration of all incoming patient data from the telemedicine project MACCS[1,18,19,21,26,39] into these EHR was enabled. For this purpose, a telemedicine dashboard was built directly in TBase, which is connected via a HL7 FHIR interface to the smartphone of the patient and the home nephrologist. Consequently, the telemedicine physician in the hospital has access to vital signs collected by the patient at home[20]. The automatic transmission of relevant patient data (including laboratory data) from the home nephrologist provides an additional clinical benefit for the physician and creates a more comprehensive database for research. In the future, with implementing new interfaces, the EHR shall be able to capture data from online patient communication (e.g., chat), IoT devices (e.g., mobile ECGs, point of care devices), and to integrate all incoming data in real-time as well as displaying patients at risk with the help of AI algorithms.

Conclusion

In summary, TBase offers a practice-approved and comprehensive electronic patient record for KTR optimized for research-oriented transplant centers. It can be easily transferred to other transplant centers due to its modularity and platform independence. Its design allows the integration of data from home monitoring and AI-driven data analyses. The TBase user group will drive its further development with the goal to foster transplantation research.

ACKNOWLEDGEMENTS:

The development of the presented EHR was supported over the last 20 years by internal research funding and public funding from different institutions and foundations.

DISCLOSURES:

The corresponding authors have nothing to declare.

REFERENCES:

1 Halleck F et al. Integrated care of renal transplant patients – Development of an electronic health care service platform. *Dialyse aktuell.* **20** (06), 285-290 (2016).

2 Sonntag, D. et al. The Clinical Data Intelligence Project. *Informatik-Spektrum.* **39** (4), 290-300, doi:10.1007/s00287-015-0913-x (2016).

3 Kara, E. et al. A Domain-adapted Dependency Parser for German Clinical Text. *Proceedings of the 14th Conference on Natural Language Processing.* (Vienna Austria KONVENS 9/2018) (2018).

4 Maier, C. et al. Experiences of Transforming a Complex Nephrologic Care and Research Database into i2b2 Using the IDRT Tools. *Journal of Healthcare Engineering.* **2019**, 5640685, doi:10.1155/2019/5640685 (2019).

5 Jensen, P.B., Jensen, L.J., Brunak, S. Mining electronic health records: towards better research applications and clinical care. *Nature Reviews Genetics.* **13** (6), 395-405, doi:10.1038/nrg3208 (2012).

6 Schmidt, D. et al. A novel tool for the identification of correlations in medical data by faceted search. *Computers in Biology and Medicine.* **85**, 98-105, doi:10.1016/j.compbiomed.2017.04.011 (2017).

7 Veit, K., Wessels, M., Deiters, W. Gesundheitsdaten und Digitalisierung – Neue Anforderungen an den Umgang mit Daten im Gesundheitswesen. *Digitale Transformation von Dienstleistungen im Gesundheitswesen VI: Impulse für die Forschung.* 19-33, doi:10.1007/978-3-658-25461-2_2 (2019).

8 Ehrler, F., Geissbuhler, A., Jimeno, A., Ruch, P. Data-poor categorization and passage retrieval for gene ontology annotation in Swiss-Prot. *BMC Bioinformatics.* **6 Suppl 1**, S23, doi:10.1186/1471-2105-6-S1-S23 (2005).

9 Esteban, C., Schmidt, D., Krompaß, D., Tresp, V. Predicting Sequences of Clinical Events by Using a Personalized Temporal Latent Embedding Model. *2015 International Conference on Healthcare Informatics.* 130-139, doi:10.1109/ICHI.2015.23 (2015).

10 Roller, R. et al. Detecting Named Entities and Relations in German Clinical Reports. *Language Technologies for the Challenges of the Digital Age.* 146-154 (2018).

11 Roller, R. et al. A fine-grained corpus annotation schema of German nephrology records. *Clinical Natural Language Processing Workshop (ClinicalNLP).* (Osaka, Japan) (2016).

12 Esteban, C., Staeck, O., Baier, S., Yang, Y., Tresp, V. Predicting Clinical Events by Combining Static and Dynamic Information Using Recurrent Neural Networks. **2016** (IEEE International Conference on Healthcare Informatics (ICHI)), 93-101, doi:10.1109/ICHI.2016.16 (2016).

13 Schmidt, D., Niemann, M., Trzebiatowski, G.L.v. The Handling of Missing Values in Medical Domains with Respect to Pattern Mining Algorithms. *Proceedings of the 24th International Workshop on Concurrency, Specification and Programming (CS&P 2015).* **1492**, 147-154 (2015).

14 Burchardt, A. Uszkoreit, H. *IT für soziale Inklusion: Digitalisierung – Künstliche Intelligenz – Zukunft für alle.* De Gruyter (2018).

15 K. Schröter. G. Lindemann, L.F. TBase2 – A Web-Based Electronic Patient Record. *Fundamenta Informaticae.* **43** (1–4), 343–353, doi:10.3233/FI-2000-43123417 (2000).

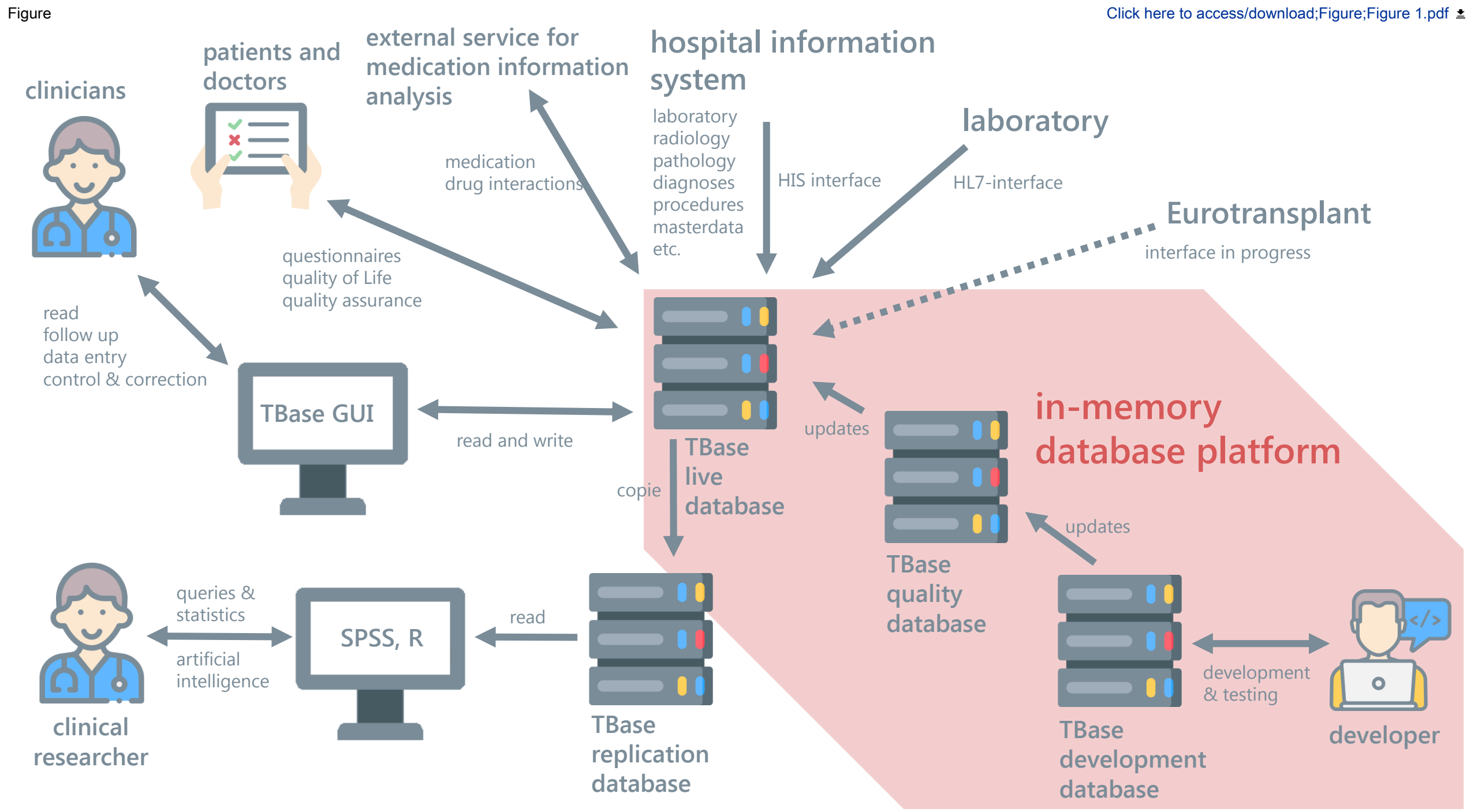
16 Duettmann, W. et al. eHealth in Transplantation. *Transplant International.* doi:10.1111/tri.13778 (2020).

17 Durr, M. et al. Late Conversion to Belatacept After Kidney Transplantation: Outcome and Prognostic Factors. *Transplantation Proceedings.* **49** (8), 1747-1756 e1741, doi:10.1016/j.transproceed.2017.05.013 (2017).
18 Halleck, F. et al. MHealth and digital management after kidney transplantation. *Nieren- und Hochdruckkrankheiten.* **46**, 474-480, doi:10.5414/NHX01883 (2017).
19 Düttmann-Rehnolt, W. et al. Neuartige Kommunikationswege und Strukturen zur Optimierung der häuslichen Versorgung am Beispiel von nierentransplantierten Patienten. *Digitale Transformation von Dienstleistungen im Gesundheitswesen VI: Impulse für die Forschung.* 407-421, doi:10.1007/978-3-658-25461-2_21 (2019).
20 S. Hils, A.S., L. Bogatyreva, D.Hauschke, P. Pisarski. Telemedical Supported Aftercare as an Innovative Project-Study Improves the Quality of Life After Living Kidney Transplantation - A Single Center Experience. *American Journal of Transplantation.* **14** (S3), 843- (2014).
21 Schmidt, D. et al. Integrierte Versorgung chronisch kranker Patienten am Beispiel von MACSS: Digitalisierung – Künstliche Intelligenz – Zukunft für alle. *IT für soziale Inklusion.* 41-50, doi:10.1515/9783110561371-005 (2018).
22 Massie, A.B., Kucirka, L.M., Segev, D.L. Big data in organ transplantation: registries and administrative claims. *American Journal of Transplantation.* **14** (8), 1723-1730, doi:10.1111/ajt.12777 (2014).
23 McAdams-DeMarco, M.A. et al. Frailty, Length of Stay, and Mortality in Kidney Transplant Recipients: A National Registry and Prospective Cohort Study. *Annals of surgery*. **266** (6), 1084-1090, doi:10.1097/SLA.0000000000002025 (2017).
24 w3schools.com. *The Most Popular Browsers*. https://www.w3schools.com/browsers/ Accessed on: 22.08.2020.
25 (HL7), H.L.S.I. *HL7 International*. https://www.hl7.org/ Accessed on: 22.08.2020, (2020).
26 Duettmann, W. et al. Digital home monitoring of patients after kidney transplantation: The MACCS platform. *Journal of Visualized Experiments.* (Accepted 2021).
27 Haas, M. et al. The Banff 2017 Kidney Meeting Report: Revised diagnostic criteria for chronic active T cell-mediated rejection, antibody-mediated rejection, and prospects for integrative endpoints for next-generation clinical trials. *American Journal of Transplantation.* **18** (2), 293-307, doi:10.1111/ajt.14625 (2018).
28 Roufosse, C. et al. A 2018 Reference Guide to the Banff Classification of Renal Allograft Pathology. *Transplantation.* **102** (11), 1795-1814, doi:10.1097/TP.0000000000002366 (2018).
29 Duerr, M. et al. Increased incidence of angioedema with ACE inhibitors in combination with mTOR inhibitors in kidney transplant recipients. *Clinical Journal of the American Society of Nephrology.* **5** (4), 703-708, doi:10.2215/CJN.07371009 (2010).
30 Lachmann, N. et al. Invited letter in response to "Predicted indirectly recognizable HLA epitopes (PIRCHE): Only the tip of the iceberg?". *American Journal of Transplantation.* **18** (2), 523-524, doi:10.1111/ajt.14535 (2018).
31 Huber, L., Naik, M., Budde, K. Desensitization of HLA-incompatible kidney recipients. *The New England Journal of Medicine.* **365** (17), 1643; author reply 1644-1645, doi:10.1056/NEJMc1109936 (2011).
32 Choi, M. et al. Low Seroprevalence of SARS-CoV-2 Antibodies during Systematic Antibody Screening and Serum Responses in Patients after COVID-19 in a German Transplant Center. *Journal of Clinical Medicine.* **9** (11), doi:10.3390/jcm9113401 (2020).

881 33 Lehner, L.J. et al. Analysis of Risk Factors and Long-Term Outcomes in Kidney Transplant
882 Patients with Identified Lymphoceles. *Journal of Clinical Medicine.* **9** (9), doi:10.3390/jcm9092841
883 (2020).
884 34 Zhang, Q. et al. The relationship between proteinuria and allograft survival in patients
885 with transplant glomerulopathy: a retrospective single-center cohort study. *Transplant*
886 *International.* doi:https://doi.org/10.1111/tri.13787 (2020).
887 35 Bissler, J. et al. Everolimus for angiomyolipoma associated with tuberous sclerosis
888 complex or sporadic lymphangioleiomyomatosis (EXIST-2): A multicentre, randomised, double-
889 blind, placebo-controlled trial. *Lancet.* **381**, doi:10.1016/S0140-6736(12)61767-X (2013).
890 36 Budde, K. et al. Everolimus-based, calcineurin-inhibitor-free regimen in recipients of de-
891 novo kidney transplants: an open-label, randomised, controlled trial. *Lancet.* **377** (9768), 837-
892 847, doi:10.1016/S0140-6736(10)62318-5 (2011).
893 37 G. Lindemann, L.F., K. Schröter, A. Schlaefer, K. Budde, H. H. Neumayer. Web-Based
894 Patient Records – The Design of TBase2. *New Aspects of High Technology in Medicine.* (2000).
895 38 BigMedilytics. *Big Data Project | BigMedilytics | The largest initiative to transform*
896 *healthcare sector*. https://www.bigmedilytics.eu/big-data-project/ Accessed on: 21.11.2020,
897 (2020).
898 39 Duettmann, W. et al. Telemedizinische Betreuung von Patienten nach
899 Nierentransplantation: Was beinhaltet MACCS? *Nieren- und Hochdruckkrankheiten.* **49** (2020).
900

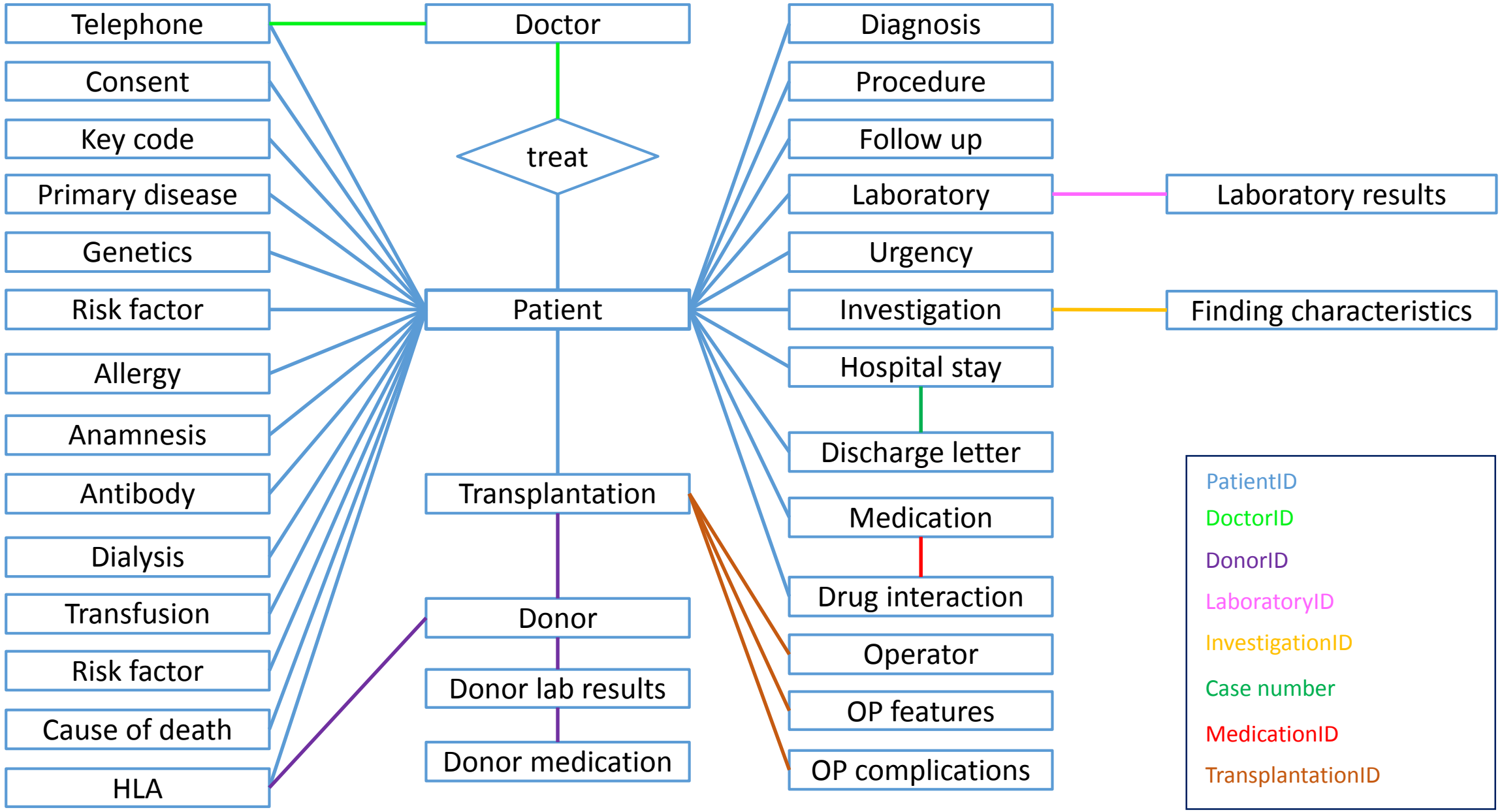
Figure

Click here to access/download;Figure;Figure 1.pdf



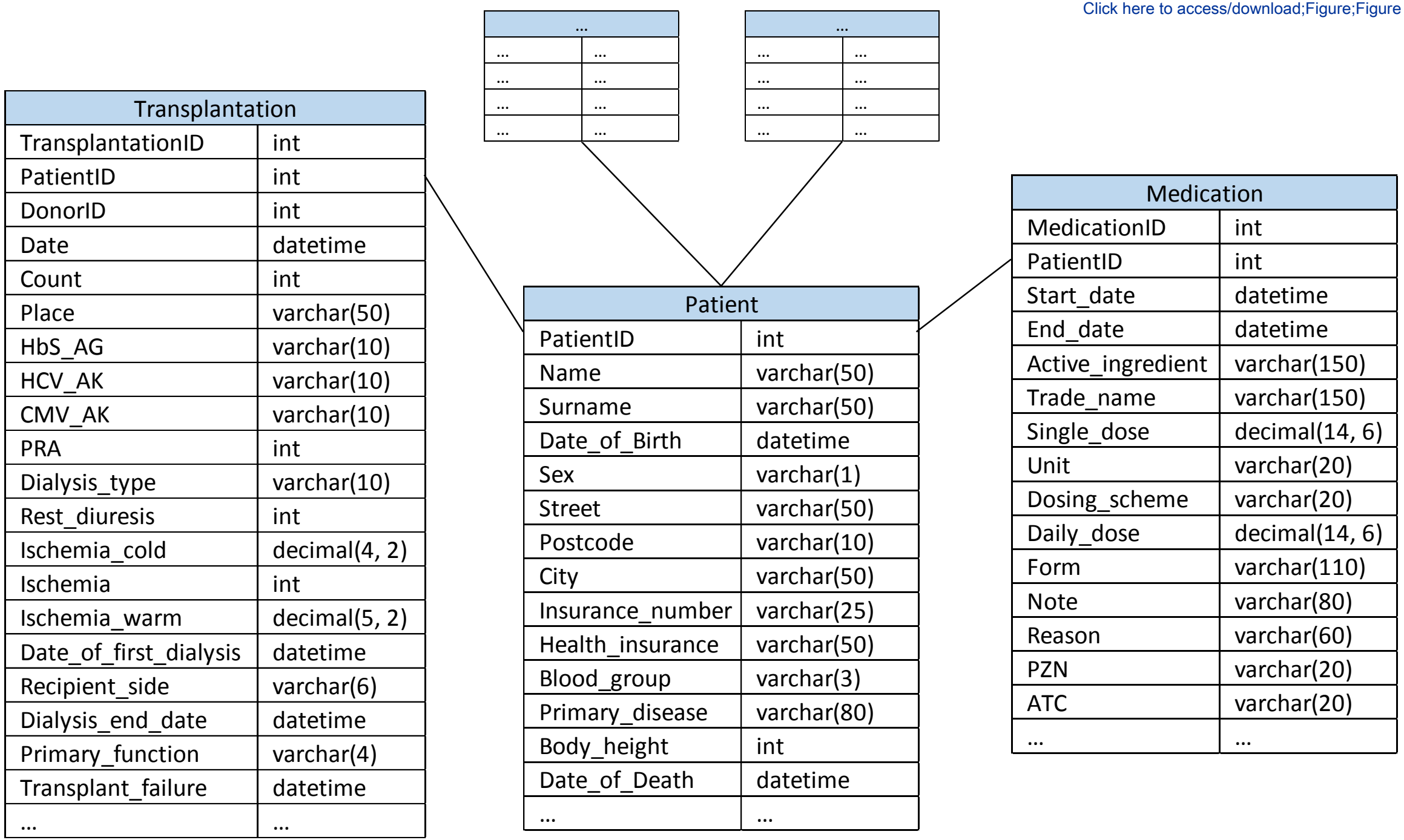
Figure

Click here to access/download;Figure;Figure 2.pdf



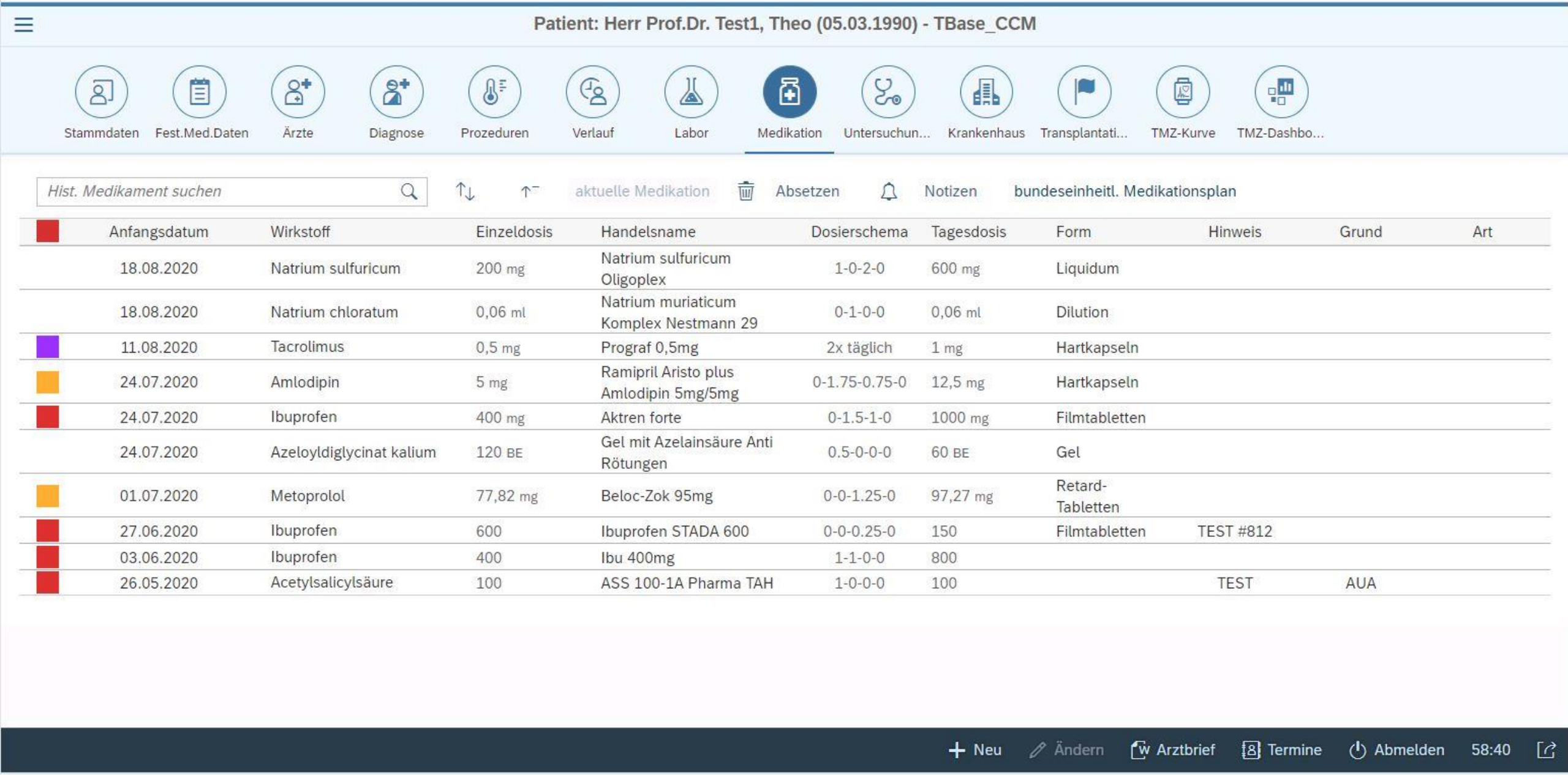
Figure

Click here to access/download;Figure;Figure 3.pdf



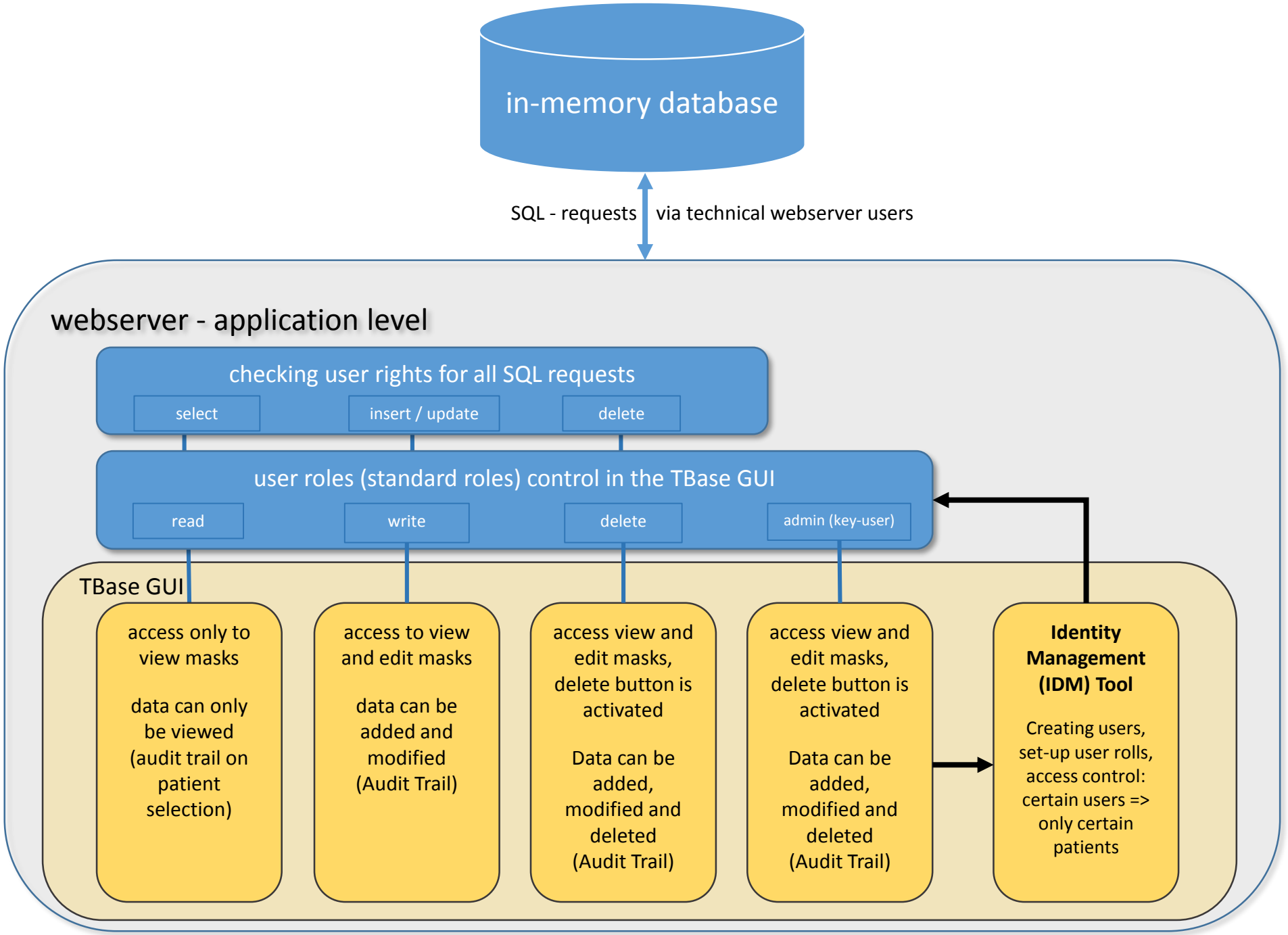
Figure

Click here to access/download;Figure;Figure 4.pdf



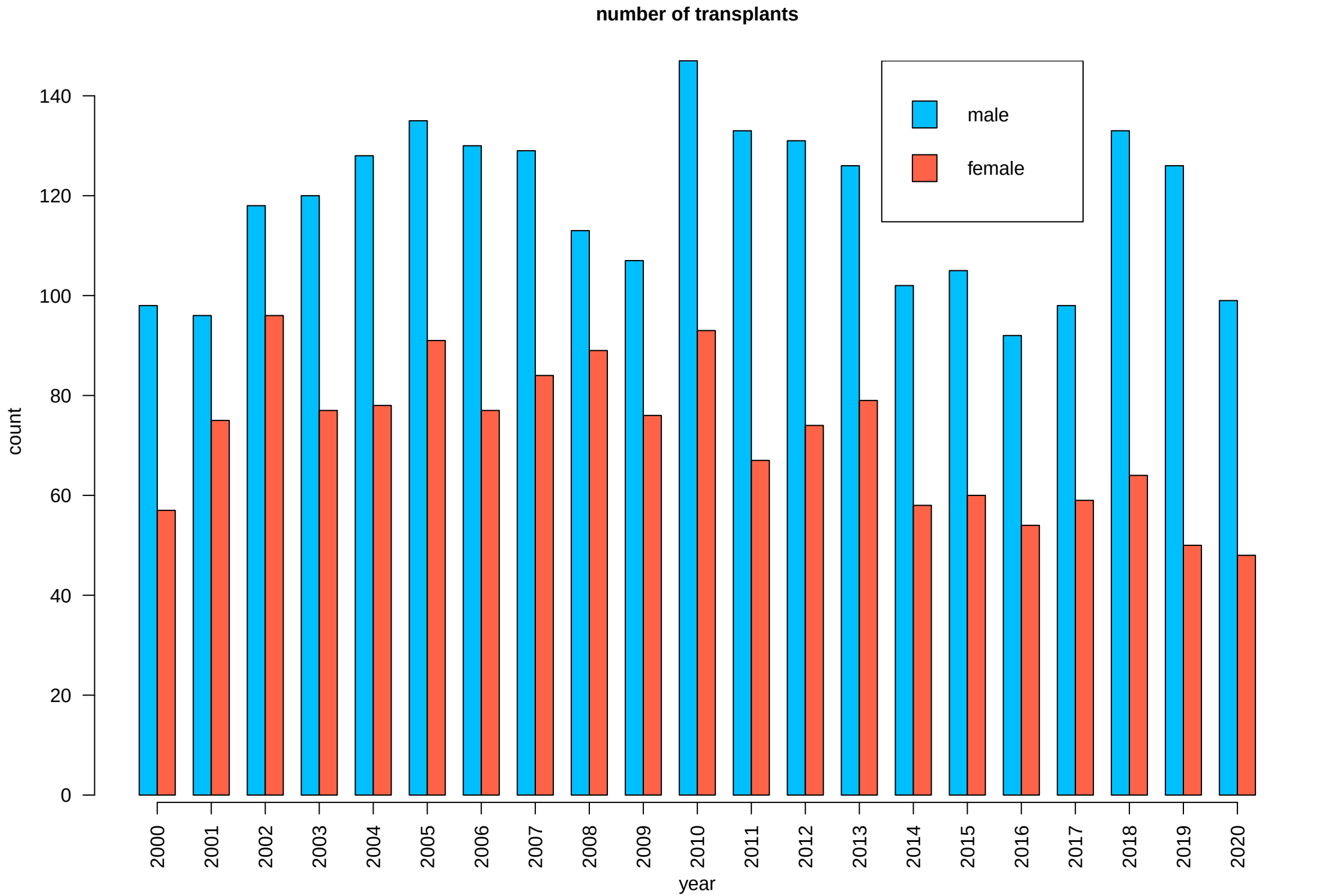
Anfangsdatum	Wirkstoff	Einzeldosis	Handelsname	Dosierschema	Tagesdosis	Form	Hinweis	Grund	Art
18.08.2020	Natrium sulfuricum	200 mg	Natrium sulfuricum Oligoplex	1-0-2-0	600 mg	Liquidum			
18.08.2020	Natrium chloratum	0,06 ml	Natrium muriaticum Komplex Nestmann 29	0-1-0-0	0,06 ml	Dilution			
11.08.2020	Tacrolimus	0,5 mg	Prograf 0,5mg	2x täglich	1 mg	Hartkapseln			
24.07.2020	Amlodipin	5 mg	Ramipril Aristo plus Amlodipin 5mg/5mg	0-1.75-0.75-0	12,5 mg	Hartkapseln			
24.07.2020	Ibuprofen	400 mg	Aktren forte	0-1.5-1-0	1000 mg	Filmtabletten			
24.07.2020	Azeloyldiglycinat kalium	120 BE	Gel mit Azelainsäure Anti Rötungen	0.5-0-0-0	60 BE	Gel			
01.07.2020	Metoprolol	77,82 mg	Beloc-Zok 95mg	0-0-1.25-0	97,27 mg	Retard-Tabletten			
27.06.2020	Ibuprofen	600	Ibuprofen STADA 600	0-0-0.25-0	150	Filmtabletten	TEST #812		
03.06.2020	Ibuprofen	400	Ibu 400mg	1-1-0-0	800				
26.05.2020	Acetylsalicylsäure	100	ASS 100-1A Pharma TAH	1-0-0-0	100		TEST	AUA	

Figure

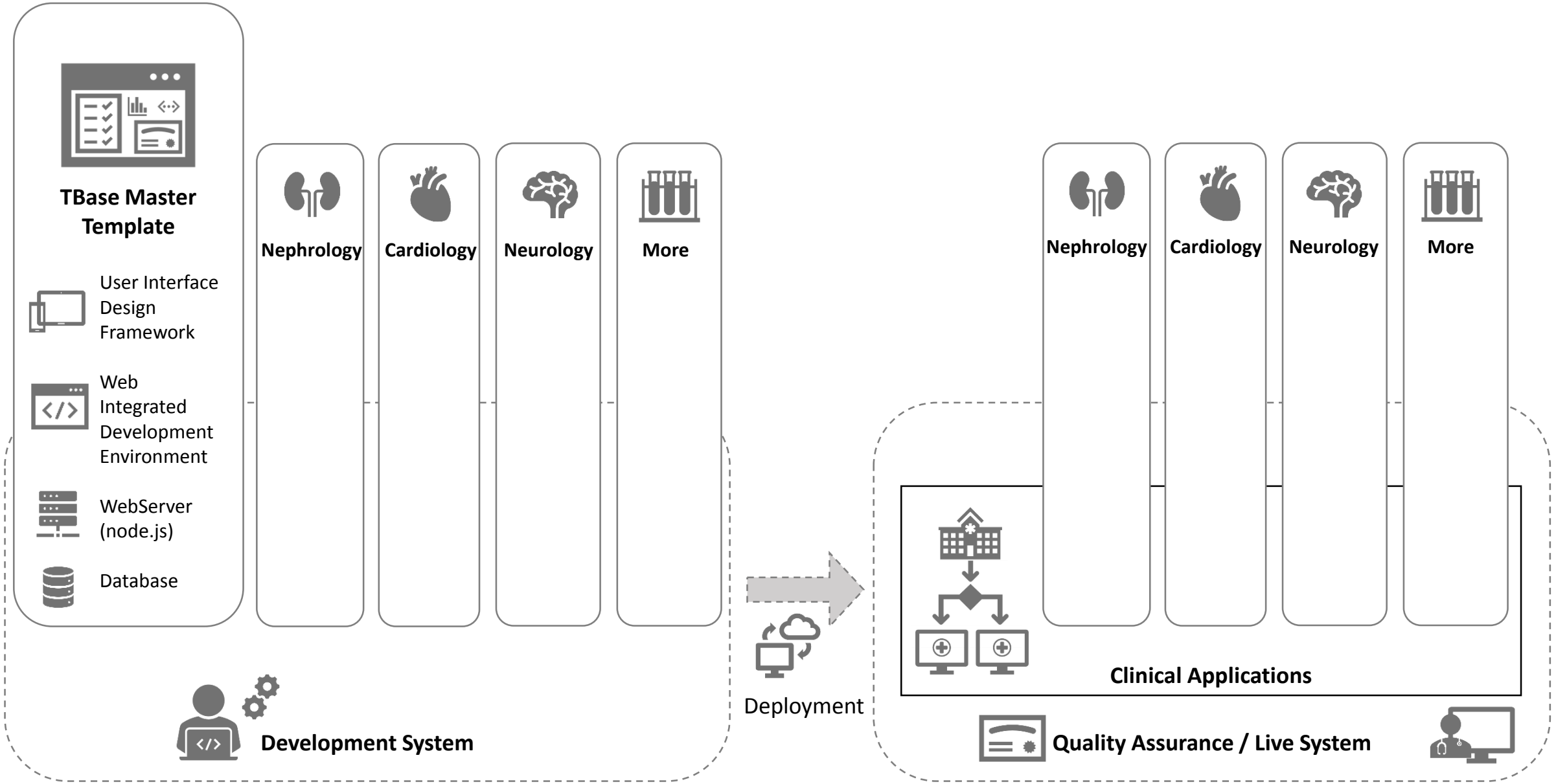


Figure

Click here to access/download;Figure;Figure 6.pdf



Figure



Disadvantages of hospital information systems (HIS)
Relevant data are not fully available
No or limited adaptation to domain-specific needs
Designed for either inpatient or outpatient care
Difficult data extraction for research purposes
Changing the HIS often results in loss of previous data.

Disadvantages of standalone research databases

Manual data entry, no automated data transfer of routine clinical data

Data entry limited to key variables, thus impeding machine learning applications

Data validation is difficult and time-consuming

Usually unemployable for new needs. Thus, a new database is created.

Single user standalone systems often result in short-lived data silos.

Benefits in clinical practice

Interfaces realize the automatic import of

- Clinical reports (laboratory, pathology, radiology, virology, microbiology)
- Hospital data
- ET donor data (in progress)

Documentation fields (partially with selection lists) allow the inclusion of numerous subject-specific data and their maintenance by the treating physicians

- Demographic data, personal data (address, telephone number, e-mail etc.), basic medical data, diagnoses, histories, procedures, laboratory data, medications, examinations, hospital data, transplantation data, donor data, rejections

More than 20 years of follow up data for individual patients through continuous development of the same database

Easy access at outpatient clinic, medical ward and emergency room to all relevant medical data including:

- Transplant data
- Examination results, laboratory values, medical reports and notes
- Medication lists (current and historical)
- Phone numbers and other core data

Semi-automated generation of

- Medical reports for external physicians
- Medication lists and laboratory values for patients

Compliance with EU GDPR data protection requirements

Graphical historical data display (e.g. multiple laboratory values over time and their correlation in a diagram with just a few clicks possible)

Display of the annual GFR decline (planned)

Monitoring dashboard for real time display of outcome parameters, waitlist, special patient groups (planned)

Facilitating quality assurance

- Automatic export of the relevant data to the hospital QA module (planned)
- Easy queries for quality control or other evaluations

Implementation of new methods in clinical routine such as telemedicine, or predictive AI models for identifying patients at-risk

Benefits for research

Easy data collection for clinical observational studies as well as for case reports / case series

Fast data access for documentation in clinical trials / studies

100% follow-up of patient and transplant survival, complete and validated data set for laboratory values, medication lists, diagnoses, procedures, and examinations, donor and transplant data available

Simple data extraction for research projects with other researchers regarding different aspects (e.g., Eurotransplant Senior Program, new drug therapies, intensive care stays, patients with postoperative complications)

High data quality due to continuous data validation in routine care by clinical users

Connection to different statistical analysis tools (e.g. SPSS, R) via standard interfaces

Easy data extraction from unstructured data (notes, letters, medical reports) by using modern text mining techniques

Analyses with artificial intelligence methods enable the development of predictive models

Analysis of new data sources from telemedicine and mobile activity trackers

Table 3

patients with kidney transplant	N = 6.317	
kidney transplant		7,595
diagnoses		220,877
procedure		332,299
laboratory		1,033,941
laboratory values		24,478,441
medications		539,922
investigations		324,339
discharge letters		54,350
donor data		6,489

Table of Materials

Name of Material/Equipment	Company	Catalog Number	Comments/Description
Developer platform SAP Web IDE	SAP SE		
GUI Toolbox SAPUI5	SAP SE		
In-memory database SAP-HANA	SAP SE		
Interface Standard HL7	Health Level Seven International		
Interface Standard HL7 FHIR	Health Level Seven International		
RStudio	RStudio Inc.		
TBase - Electronic Health Record	Charité - Universitätsmedizin Berlin		
Webserver SAP-HANA XSA	SAP SE		

Response Letter

Dear editor, dear reviewers,

thank you very much for your valuable feedback, which helped us to further improve our manuscript. In the following, we will explain point by point how we addressed the editorial and reviewers comments. We are confident that our manuscript now is acceptable for publication in your journal.

Danilo Schmidt on behalf of the co-authors

Editorial comments:

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues.

Answer: We proofread the manuscript twice before resubmission.

2. Please define the abbreviations before use. (EHR, API, HL7, EU GDPR, ICD-10, DSN, ODBC, JDBC)

Answer: Sorry for this mistake, in the revised manuscript we now defined abbreviations before use.

3. Please ensure that, for in-text formatting, the corresponding reference numbers appear as numbered subscripts after the appropriate statements. Multiple references must be separated by commas.

Answer: In the revised manuscript we edited the references according the guidelines.

4. The Protocol should be made up almost entirely of discrete steps without large paragraphs of text between sections. Please simplify the Protocol so that individual steps contain only 2-3 actions per step and a maximum of 4 sentences per step.

Answer: We revised the protocol accordingly and checked that now all protocol steps meet these criteria.

5. In the JoVE Protocol format, “Notes” should be concise and used sparingly. They should only be used to provide extraneous details, optional steps, or recommendations that are not critical to a step. Any text that provides details about how to perform a particular step should either be included in the step itself or added as a sub-step. Please consider moving some of the notes about the protocol to the discussion section.

Answer: We reduced the amount of “Notes” in the revised protocol. In order to discuss some aspects in greater detail, we expanded the discussion about pathology classifications in Line 682 – 687 (see next comment below) and medication data to lines 611 – 622.

6. Line 326: Please add the details regarding the basis of the classification. An appropriate citation would suffice.

Answer: We added information about the classification of kidney transplant biopsies in Line 330 and in Line 682-687 in the discussion part together with an appropriate citation.

7. Line 358: Please adjust the numbering of the Protocol to follow the JoVE Instructions for Authors. ("2.8.3" should be "2.9.3", "2.8.4" should be "2.9.4")

Answer: We changed the numbering according your instructions.

8. Each Figure Legend should include a title and a short description of the data presented in the Figure and relevant symbols. The Discussion of the Figures should be placed in the Representative Results. Details of the methodology should not be in the Figure Legends, but rather the Protocol.

Answer: We shortened the description of Table 1 and 2.

9. Please do not use any abbreviations for the journal titles and book titles.

Answer: We replaced the abbreviations for journal titles by full titles.

10. Figure 1: Please remove the commercial branding: SAP, SAP HANA (also in Figure 5), Labor Berlin, etc.

Answer: We edited the figures accordingly and removed all commercial brandings. In the text we only mention the commercial product once, because it is important for the reader to know the exact current development system of the database.

11. Figure 6: Please define the dataset labels on the top right corner so that it does not cove the represented data bars.

Answer: The dataset was recreated in a way that the label does not cover the data bars.

12. Please move the code from step 7.2 to a supplementary file.

Answer: The code from step 7.2. was moved to a supplementary file.

13. Please include an Acknowledgements section, containing any acknowledgments and all funding sources for this work.

Answer: An Acknowledgements section as well as an Authors' contributions section was added.

14. Please spell out the journal titles.

Answer: The journal titles were spelled out in the references.

Reviewer #1:
Major Concerns:
"There is not a real description of the relationships and transformations between the EHR and the TBase database structures. This needs to be much more explained. How have the authors mapped the elements of the EHR to the working structure of the database? Moreover there is nor any description of TBase itself. How is it organized? Why is it such a good database for secondary use? Why is it so smoothly connected to the EHR. An example or figure could be very helpful. The authors speak a lot about its operation but almost nothing about its working structure and characteristics."

Answer: We explicated throughout the revised manuscript that TBase and the electronic health record (EHR) are the same. The database and the graphical user interface (GUI) together, build the EHR for kidney transplant recipients named TBase. We added Figure 3, showing exemplarily which data types are used for representation of different clinical data in the underlying database. In line 115-121 we

explained briefly, in which way the database structure is mapped to the graphical user interface of TBase and that this is the electronic health record, which end users like physicians and nurses use.

Minor Concerns:
“On page 3 the "concept of data validation in clinical routine" needs to be explained and clarified.”

Answer: This passage found in the Abstract was replaced by “This is achieved by the concept of data validation in clinical routine in which clinical users and patients have to rely on correct data for treatment and medication plans and thereby validate and correct the clinical data in their daily practice” in lines 51-54. Similarly, we added a sentence on page 3 (Line 130-134).

“I'm not sure if Figure 4 on p. 10 refers to the the text in that page according to the foot label of that figure on p. 11.”

Answer: We corrected the text of all figures.

“Where is Figure 5 referenced in text?”

Answer: We referenced Figure 5 in section “Management and Data Protection” in line 183 in the new manuscript.

Reviewer #2:
Major Concerns:
T-Base was first released in 1999. A major concern is that the data quality has been changing dynamically during the data period. This manuscript described data quality in general, but there is no discussion on how the data quality has been changed during the data period and how those changes affected data integration.

Answer: We commented how changing data quality over time affects the data recorded in our database and how we addressed the problems arising. We therefore included lines 665 to 704 into our revised discussion. We herein explained that laboratory, medication, procedure and diagnosis data did not change remarkably over time. For pathology data, we reclassified all kidney transplant biopsies according to the latest BANFF 2017 classification. For other data such as medical reports or examination reports, which are currently recorded as strings, we aim to use methods from Natural Language Processing to structure those data in the future. As a limitation we stated that some data, such as those about genetic diseases cannot be generated retrospectively for previously transplanted patients.

Minor Concerns:
One interesting forward-looking expansion is the implementation of Artificial Intelligence, but there is no enough detail about this potential expansion

Answer: We acknowledge the demand for more detail about the implementation of methods from Artificial Intelligence into our EHR, but have to mention that most of our work in this field is preliminary. In the previous version of our manuscript, we already outlined the design of a study that is still ongoing, in which we are comparing the predictive accuracy of an AI-based model with experienced human physicians. We added information about the underlying AI method, that we use “Gradient Boosted Regression Trees” for our current model in line 735. We additionally envisioned the use of Natural Language Processing to structure text based data from medical reports or examination reports in lines 740 to 744. These data may further increase the predictive accuracy of the previously mentioned AI based prediction models.

Click here to access/download

Supplemental Coding Files

Code 1.docx