

THE STRUCTURED STORAGE OF ONCOLOGICAL CHEMOTHERAPEUTIC REGIMENS

Contribution to Standardization of Therapeutic Procedures in Current Oncology

D. Klimes¹, L. Dusek¹, M. Kubasek¹, J. Novotny², J. Finek³ and R. Vyzula⁴

¹ *Institute of Biostatistics and Analyzes, Masaryk University, Kamenice 126/3, Brno, Czech Republik*

² *Department of Oncology, General Teaching Hospital, Prague, Czech Republic*

³ *Charles University Prague, Medical Faculty Pilsen, Pilsen, Czech Republic*

⁴ *Masaryk Memorial Cancer Institute, Brno, Czech Republic*

Keywords: Cancer Chemotherapy Protocols, Antineoplastic Combined Chemotherapy Regimens, XML, Clinical Database.

Abstract: The aim of the chemotherapeutic regimens (CHR) digitalization project is the proposal of a universal structure and creation of a publicly accessible database of contemporary CHR as a universal utility for the communication and evaluation of contemporary and newly defined clinical schedules in anti-tumor chemotherapy. After analysis of contemporary anti tumor CHR a standard XML structure was proposed, which enables the recording of simple CHR from the field of chemotherapy in solid adult tumors, and also has the potential of recording the complex treatment protocols in the field of paediatric oncology. The resulting XML documents were saved on a web server. A publicly accessible CHR database was constructed. There were a total of 130 XML documents with definitions of individual CHR in the first phase. Linked to this data store, three examples of web applications were added to demonstrate the potential uses of this newly created database.

1 INTRODUCTION

Chemotherapy, along with surgery and radiotherapy is an irreplaceable part of the clinical treatment of oncological illnesses across nearly all diagnoses. To be more exact, we can define the term of chemotherapy as encountered in this article. It is the administration of individual preparations or combinations of preparations from the anatomical-therapeutic-chemical group (ATC) L01, which are anti-tumor preparations. This ATC group currently consists of nearly a hundred generic preparations, which are divided into 5 basic groups: Alkylating agents (L01A), Antimetabolites (L01B), Plant alkaloids and other natural products (L01C), Cytotoxic antibiotics and related substances (L01D) and Other antineoplastic agents (L01X). Whether these preparations are in practice applied either separately in monotherapy or in combination, we refer to them as a chemotherapeutic regimen (CHR). The more complicated plans for the application of cytostatics are classified in paediatric oncology as treatment protocols. Standard CHR are based on

clinical studies and are published in the oncological journals (Goldberg et al., 2004) (Henderson et al., 2003) (Citron et al., 2003) and subsequently in national or international guidelines (NCCN, 2006) (CLS JEP, 2005).

However, CHR are not merely a list of applied cytostatics in administered doses; the definition of a CHR has its own basic rules, which are however not strictly defined anywhere. The main features of CHR include: doses of cytostatics are defined most often according to the surface area of the patient's body or their weight. CHR are applied in cycles, i.e. defined time segments of the treatment are repeated several times. There can be one or more repeated segments. The days of application are identified relative to the first day of each cycle (example Cycle 1- Day 1, Cycle 1- Day 8). These features are in practice routinely used (see Figure 1).

Klimes D., Dusek L., Kubasek M., Novotny J., Finek J. and Vyzula R. (2008).

THE STRUCTURED STORAGE OF ONCOLOGICAL CHEMOTHERAPEUTIC REGIMENS - Contribution to Standardization of Therapeutic Procedures in Current Oncology.

In *Proceedings of the First International Conference on Health Informatics*, pages 148-154

Copyright © SciTePress

FEC chemotherapy
Cyclophosphamide 75 mg/m PO days 1-14
Epirubicin 60 mg/m IV days 1 & 8
5-Fluorouracil 500 mg/m IV days 1 & 8
With cotrimoxazole support.
Cycled every 28 days for 6 cycles.

Figure 1: An example of CHR clinical definition (NCCN Clinical Practice Guidelines in Oncology™ ©2006).

The aforementioned structure is frequently used, however not standardized, as to prevent a wider use of information technology. The vendors of a hospital information system can create a special application module for chemotherapy with a specific, internal data structure or chemotherapy data is stored only as a sequence of applications of cytostatics without specific details. This situation hinders on the one hand fast electronic transmission of new CHR and the general intercommunication between computer applications in this field, and on the other hand the more advanced use of technology, for instance in the field of the assessment and adherence to standard CHR in clinical practice. A possible solution is the creation of a structured data store of current CHR, which will mirror current clinical guidelines and which will be freely accessible to both clinics and HIS providers. This article describes the proposed structure of such a data store and the experience gained during its construction. The practical usage of such a database of CHR is demonstrated with the example of a web portal, whose components are web applications which use the CHR database as its datasource.

2 METHODS

Preliminary analysis of problems showed that the construction of a structured database of CHR calls for certain steps:

- 1) The proposal of the structure of CHR
- 2) The unique identification of standard CHR
- 3) The digitalisation of current standard CHR

CHR can be considered as a type of structured document, which is why the XML language was chosen for its recording. Standard XML offers a workable computational structure. In addition, it also provides tools for the internal validation of structures (XML schema) and tools for transforming documents into different, more user-friendly formats, for example HTML pages (XSLT). The proposed XML structure of elements and attributes

was developed dynamically through the analysis of clinical definitions of standard CHR for individual oncological diagnoses. The source of the definitions was the National and International oncological guidelines (NCCN, 2006) (CLS JEP, 2005) and the internal source of Masaryk memorial Cancer Institut, a specialised hospital for the treatment of oncological diseases in the Czech Republic. The template for CHR was modelled using XML schemes, which enabled the definition of individual elements and attributes. The mandatory/optional properties and frequency of repetition of elements are defined by the maxOccurs and minOccurs indicators.

2.1 Header of CHR

The created XML structure of CHR consists of two parts. The first deals with the identification of CHR and the possibility of its use in oncological diagnosis (header). An example of this header is shown in figure 2.

```
<name>AC(Fisher)</name>  
<sysname>(1;60.0;mg/m2;iv)A+(1;600.0;mg/m2;iv)C&#21  
</sysname>  
<diagnosis>  
  <ICD10>C50</ICD10>  
  <line>1</line>  
  <purpose>adjuvant</purpose>  
</diagnosis>
```

Figure 2: An example of CHR header definition.

The header contains the element name which was used for the clinical identification of CHR. This name was adopted either directly from clinical guidelines or from clinical identification in the information system of MMCI. These clinical names do not guarantee uniqueness and do not adhere to any strict rules. On the other hand the element sysname was added to the individual definitions of CHR on the basis of their internal structure. The detailed principles of their creation are described below.

The diagnosis element lists the individual oncological diagnoses that each individual CHR can treat. This is a complex element, which contains the following nested elements. The code for diagnoses is introduced according to the international classification ICD-10 in same named nested elements. The element line defines the CHR which is suitable (or approved) for specific lines of treatment.

The element purpose specifies whether a CHR is designed for adjuvant or palliative treatment. The elements line and purpose can, within the one element diagnosis, occur repeatedly in cases where the CHR is intended for more lines or for both basic purposes of treatment. In cases where the CHR is used for more diagnoses, the whole complex element Diagnosis is repeated.

2.2 Body of CHR

The second part of the structure describes the administration of given CHR (body). A standard CHR was divided into the following components

- The name of administered cytostatics
- Dosage of individual cytostatics
- Units of doses
- Method of administration
- Day relative to cycle of administration of cytostatic
- Duration of one cycle in days
- Total of completed cycles

Primarily proposed scheme for the body of the CHR is presented in figure 3.

```

<group>
<id_group>1</id_group>
<interval>21</interval>
<noc>4</noc>
<drug>
.....
<drug>
<drug>
.....
<drug>
</group>
<group>
<id_group>2</id_group>
.....
</group>

```

Figure 3: The main frame of CHR body definition.

The element *interval* indicates the duration of one cycle of chemotherapy in days. The length of one cycle can be defined as the number of days between day D1 on one cycle and D1 on the following cycle. The actual length of the last cycle can only be determined from the last defined day of administration and cannot be compared with the length of the preceding cycles.

The element *noc* (number of cycles) shows the total of applied cycles. This parameter is limited by

clinical guidelines to a small number of CHR. These guidelines often merely provide a recommendation for the repetition of cycles. In practice the number of applied cycles is decided by the actual state of health of the patient. In cases where details were not explicitly known, the value of this element was set to 0.

The complex element *drug* describes the application of individual cytostatics within the framework of one cycle of chemotherapy. Since many cytostatics are applied in CHR, the element *drug* is referred to as a recurrent element. The element *drug* encapsulates the nested elements for labelling cytostatics, their dosage, day of administration and method of administration. As soon as we try to identify individually applied cytostatics, the problem of their individual classification arises. Existing practice is to cite the full generic name of the cytostatic, which in certain cases involves the brand name of the medication. The ATC codes of identification of cytostatics are not used in clinical practice, however they are ideal for computer processing. We therefore decided to include in the XML structure both an element for the generic name of cytostatics (*name*) and an element for the ATC code (*ATC*). The element *name* can be inserted for each cytostatic repeatedly with the attribute *lang* for various language versions of classification. For the purpose of the systematic naming of CHR (described below) the element *abbr* was also defined for the abbreviated names of cytostatics.

Doses of cytostatics, units of dosage of cytostatics, method of administration and relative day of administration were included within the complex element *administration*, whose title was shortened to *adm* for practical reasons. The element *dose* was defined for the dosage of cytostatics as a real number, for units of dosage the enumerative element *unit*, for the method of administering the enumerative element *mode* and for the day of administering the whole number element *unit*, respectively the two whole number elements *start_day* and *end_day*.

In anti tumour chemotherapy, the dosage of cytostatics is most commonly defined by the calculation of the surface area of the patient or by their weight. For this reason the element *unit* was defined with the values mg/m^2 and mg/kg . Carboplatin has special dosage, where the dose is defined in AUC. The resulting dose of this drug is calculated depending on the laboratory value parameters creatinine clearance (CrCl) according to the formula:

dose(mg) = target AUC (mg/ml/min) * (CrCl +25) (ml/min)

Two basic methods of administering cytostatics (modes) are used, per oral and intravenous administration. More detailed categorization can be considered, for instance distinguishing between intravenous administrations according to the length of infusion. In the basic proposal there is the difference between bolus administration and infusion. For example in the case of the regimens FOLFOX 4 and FOLFOX 6, which are used in the treatment of colorectal carcinoma, it is necessary to differentiate between the bolus dosage of fluorouracil from its subsequent longterm infusion. The element *mode* can thus include the values *iv*, *iv-bolus* and *po*.

Day of administration is in clinical practice presented as Dx, where x is the consecutive number of days from the administration of the first preparation in a cycle. Individual cytostatics can be repeatedly administered within a cycle, either on chosen days (e.g D1, D8) or daily in the course of an appointed time period (e.g D1-D14). For the first variant, the element *day* can be repeated within the complex element *adm*, and in the second variant there are, in place of the element *day*, two elements *start day* and *end day*. An example of the complex element *drug* is illustrated in figure 4.

```
<drug>
<name lang="cz">Cyklofosamid
</name>
<name lang="eng">Cyclophosphamide
</name>
<atc>L01AA01</atc>
<abbr>C</abbr>
<adm>
<day>1</day>
<dose>600</dose>
<unit>mg/m2</unit>
<mode>iv</mode>
</adm>
</drug>
```

Figure 4: An example of the complex element *drug*.

For 'multigroup' regimens such AC+paclitaxel (see figure 5) it was necessary to extend the presented concept.

Doxorubicin 60 mg/m IV day 1
 Cyclophosphamide 600 mg/m IV day 1
 Cycled every 21 days for 4 cycles.
 Followed by
 Paclitaxel 175-225 mg/m by 3 h IV infusion
 day 1
 Cycled every 21 days for 4 cycles.

Figure 5: An example of multigroup CHR (NCCN Clinical Practice Guidelines in Oncology™ ©2006).

For this type of CHR another encapsulating structure was added to the XML scheme. This took the form of the specific complex element *group*, which contains all the defined elements of the body of the CHR and can be repeated. For the identification of groups, the element *id_group* was added which contains the consecutive number of the group. For multigroup CHR it is necessary that the element *noc* must not be zero, at least for each group except the last.

2.3 Systematic Naming of CHR

To prevent duplication in the database of CHR, a concept was sought after, which would ensure the individual identification of each of the stored CHR. The identification of CHR used in clinical practice seemed unsuitable, because as often happens, one CHR has more than one name, or one name refers to more than one CHR.

A unique standard naming of CHR was inspired by Logical Observation Identifiers Names and Codes (LOINC). LOINC is a structured classification of laboratory methods (Huff et al., 1998). There is a systematic name for each item (in this example the laboratory method) consisting of individual components which the method uniquely refers to. In the case of LOINC, the names of laboratory methods include the components, property, timing, system precision and method. Similarly, it is possible to create a unique identification system of CHR. The following requirements were necessary to be taken into account within the proposals for the rules for the systematic creation of naming of CHR:

- The naming has to be unique
- The naming must be automatically generated from definitions of the CHR
- All key components of the scheme must be coded into the name
- The name must remain "human readable"

The schematic name is created according to the following syntactic rules:

- Administered drugs are classified with a unique abbreviation
- Drugs are alphabetically sequenced according to abbreviations and are divided by the symbol plus (+)
- After the listing of the drug the duration of cycle in days is added after the symbol(&)
- For every drug, the following items are defined in round brackets separated by a semicolon (;)

The first entry in the brackets is the day of administration. It can be in the form of a number (1), a list of numbers separated by commas (1, 8) or an interval (1-14). The second entry is the dosage of the medication. The third entry is the abbreviated name for the method of administration. The fourth entry is the unit of dosage.

Abbreviations used for cytostatic medications are summarized in Table 1.

Abbreviations used are parts of the proposal for the standardization of structures of CHR, because currently the standard for the abbreviated identification of cytostatics has not been found (on the webpages of NCI (NCI Drug Dictionary National Cancer Institute, 2005) only recommended abbreviations can be found and all the well known synonyms for given preparations). Clinically established abbreviations are often diagnosis-specific, for example cisplatin is listed under the letter P in CHR such as BIP or BEP, in the regimen M-VAC it is classified under the letter C, which is however in the majority of cases used as the abbreviation for cyclophosphamide. Due to the fact that the number of well known cytostatics is very similar to the number of existing chemical elements for which two symbols are sufficient for the abbreviated symbols, a similar concept was used for the identification of cytostatics. The names of the most frequently used cytostatics are written in the table, each is recorded with a NCI abbreviation, ATC code and proposed two letter identification, which issues from the generic name of the cytostatic.

For multigroup CHR the concept was further developed with square brackets enclosing individual groups. The number of cycles is indicated before the brackets separated by asterisks (*). Groups are separated with the symbol +. An example of the identification of the CHR AC+paclitaxel is illustrated in figure 6.

Table 1: Summary of most frequently used cytostatic agents.

Cytostatic agent	ATC - code	NCI abbreviation*	Used abbreviation
bevacizumab	L01XC07	?	Be
bleomycin	L01DC01	BLEO	B
busulfan	L01AB01	BU, BUS	Bu
capecitabine	L01BC06	CAPE	Ca
carboplatin	L01XA02	CBDCA	Cb
carmustine	L01AD01	BCNU	Be
cetuximab	L01XC06	MOAB C225	Ce
chlorambucil	L01AA02	CHL, CLB	Cl
cisplatin	L01XA01	CDDP	P
cyclophosphamide	L01AA01	CTX	C
cytarabine	L01BC01	ARA-C	Cy
dactinomycin	L01DA01	DACT	Ac
dacarbazine	L01AX04	DTIC	De
daunorubicin	L01DB02	DNR	Dn
docetaxel	L01CD02	TXT	Dt
doxorubicin	L01DB01	DOX	A
epirubicin	L01DB03	EPI	E
erlotinib	L01XX34	OSI 774	Er
estramustine	L01XX11	EM	Em
etoposide	L01CB01	VP-16	Et
fludarabine	L01BB05	FAMP	Fl
fluoruracil	L01BC02	5-FU	F
gefitinib	L01XX31	ZD 1839	Ge
gemcitabin	L01BC05	dFdC	G
ifosfamide	L01AA06	IFF, IFO	If
irinotecan	L01XX19	CPT-11	I
melfalan	L01AA03	L-PAM	MI
methotrexate	L01BA01	MTX	M
mitomycin	L01DC03	MITO	Mi
mitoxantrone	L01DB07	DHAD	Mx
oxaliplatin	L01XA03	1-OHP, L-OHP	Oh
paclitaxel	L01CD01	TAX	Ta
pemetrexed	L01BA04	LY231514	Pe
prednimustine	L01AA08	?	Pr
procarbazine	L01XB01	PCB	Pc
raltitrexed	L01BA03	?	Ra
rituximab	L01XC02	MOAB IDEC-C2B8	Ri
temozolomide	L01AX03	TMZ	Tm
thiotepa	L01AC01	TSPA	Ts
topotecan	L01XX17	TOPO	To
trastuzumab	L01XC03	MOAB HER2	Tr
vinblastine	L01CA01	VBL	V
vincristine	L01CA02	VCR	Vc
vinorelbine	L01CA04	VNB	Vn

* NCI Drug Dictionary (NCI Drug Dictionary National Cancer Institute, 2005)

4*[(1;60.0;mg/m2;iv)A+(1;600.0;mg/m2;iv)C&21]+
4*[(1;175.0;mg/m2;iv)Pt&;21]

Figure 6: An example of the identification of the CHR AC+ paclitaxel

This identification of CHR is stored in a XML file in the element *sysname*. The element *sysname* acts as a unique identifier, a type of "fingerprint" of

the CHR and its primary function is to prevent duplication in the data system.

3 RESULTS

A total of 160 CHR were entered into the database. The definitions were formed according to the Czech national guidelines for the cytostatic treatment of solid tumors (CLS JEP, 2005). The validity of entries according to International standards was verified with reference to the International guidelines NCCN (NCCN, 2006).

3.1 CHR Library Applications

As a demonstration of the use of structured records of definitions of CHR a publicly accessible web application was developed with three basic functions: The Central Library of Chemotherapeutic Regimens, Dose Intensity (DI) Calculator and Therapy Organiser.

3.1.1 Search Engine

The Central Library is a simple search engine, that, according to user entered criteria, searches and displays the definition of the corresponding regimen. Registered users have the possibility to add textual commentaries to each CHR with supplementary information, while for non registered users all information is presented in a read only format.

3.1.2 Dose-Intensity Calculator

The DI Calculator enables users to calculate the dose-intensity for selected CHR according to the methods in (Hryniuk et al., 1984), and to compare this with the actual intensities of cytostatics administered to the patient in question.

3.1.3 Therapy Organizer

The Therapy Organizer enables users to devise time plans for the administration of chosen CHR. It is possible to display and print this plan in the form of a calendar with suggested days and dosages for individual cytostatics. The functions mentioned are interconnected, for example, search results from the Central Library of CHR can be directly used to create a time plan in Therapy Organizer.

The applications are accessible at the internet address
<http://dios.registry.cz/?sec=software&lang=en>

3.2 CHR Derivation

Thanks to a structured CHR library it is possible to derive standard regimen only from a list of applied drugs and dates of administration. This is useful when dose-intensity is evaluated and available data doesn't contain name of standard regimen. During pilot tests there was success in correctly deriving 98% of initial CHR from 180 patients who had been administered chemotherapy for breast carcinoma.

4 CONCLUSIONS

CHR and their administration are routine practice in contemporary oncology. The development of a structured, electronic database of standard CHR can help the faster propagation of information about new CHR and at the same time enable assessment of their adherence in clinical practice. The database is created from XML documents, where every file represents one CHR.

Unlike other printed or electronic sources about CHR, this database contains only clear, structured records of regimens. These records are inserted in cooperation with expert oncologists. The result is a new, always up-to-date information source that forms the base for Dose Intensity Analysis and also can be used in other computer applications in anti-tumour therapy area.

ACKNOWLEDGEMENTS

Project DIOS, which addresses problems in anti tumour chemotherapy, is supported by the grant 2608 from Ministry of Education of the Czech Republic and by the Amgen Inc.

REFERENCES

- Goldberg, R. M., D. J. Sargent, R. F. Morton, C. S. Fuchs, R. K. Ramanathan, S. K. Williamson, B. P. Findlay, H. C. Pitot, S. R. Alberts, 2004. A Randomized controlled trial of fluorouracil plus leucovorin, irinotecan, and oxaliplatin combinations in patients with previously untreated metastatic colorectal cancer, *Journal of Clinical Oncology* 22, no. 1: 23-30.
- Henderson, I. C., D. A. Berry, G. D. Demetri, C. T. Cirincione, L. J. Goldstein, S. Martino, J. N. Ingle, M. R. Cooper, D. F. Hayes, K. H. Tkaczuk, G. Fleming, J. F. Holland, D. B. Duggan, J. T. Carpenter, E. Frei, R.

- L. Schilsky, W. C. Wood, H. B. Muss, L. Norton, 2003. Improved outcomes from adding sequential paclitaxel but not from escalating doxorubicin dose in an adjuvant chemotherapy regimen for patients with node-positive primary breast cancer, *Journal of Clinical Oncology* 21, no. 6 (2003): 976-83.
- Citron, M. L., D. A. Berry, C. Cirincione, C. Hudis, E. P. Winer, W. J. Gradishar, N. E. Davidson, S. Martino, R. Livingston, J. N. Ingle, E. A. Perez, J. Carpenter, D. Hurd, J. F. Holland, B. L. Smith, C. I. Sartor, E. H. Leung, J. Abrams, R. L. Schilsky, H. B. Muss, L. Norton, 2003. Randomized trial of dose-dense versus conventionally scheduled and sequential versus concurrent combination chemotherapy as postoperative adjuvant treatment of node-positive primary breast cancer: First report of intergroup trial C9741/cancer and leukemia group B trial 9741, *Journal of Clinical Oncology* 21, no. 8, p. 1431-39.
- NCCN Clinical Practice Guidelines in Oncology™ NCCN, (accessed 1.10.2006). Available from http://www.nccn.org/professionals/physician_gls/f_guidelines.asp.
- Zasady cytostaticke lecby malignich onkologickych onemocneni: Ceska onkologicka spolecnost CLS JEP, 2005.
- Huff, S. M., R. A. Rocha, C. J. McDonald, G. J. De Moor, T. Fiers, W. D. Bidgood, Jr., A. W. Forrey, W. G. Francis, W. R. Tracy, D. Leavelle, F. Stalling, B. Griffin, P. Maloney, D. Leland, L. Charles, K. Hutchins, J. Baenziger, 1998. Development of the Logical Observation Identifier Names and Codes (LOINC) vocabulary, *Journal of the American Medical Informatics Association* 5, no. 3, p. 276-92.
- NCI Drug Dictionary National Cancer Institute, 2005 (accessed 12.9. 2005). Available from <http://www.cancer.gov/drugdictionary/>.
- Hryniuk, W., H. Bush, 1984. The Importance of Dose Intensity in Chemotherapy of Metastatic Breast-Cancer, *Journal of Clinical Oncology* 2, no. 11, p. 1281-88.