THE STRUCTURED STORAGE OF ONCOLOGICAL CHEMOTERAPEUTIC REGIMENS

(version 1.0)

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

Figure 2: An example of CHR header definition

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

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 HIS of MOU. 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 designated for adjuvant or palliative treatment. The elements *line* and *purpose* can, within the one element *diagnosis,* occur repeatedly in examples 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.

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.

Figure 3: The main frame of CHR body definition

```
<interval>21</interval>
<noc>4</noc>
<drug>
        .......
<drug>
<drug>
        .......
<drug>
```

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 repetion 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 dosages, 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* is possible to 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, for practical reasons was shortened to *adm.* 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 categorisation 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.

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

```
<drug>
        <name lang="cz">Cyklofosfamid</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>
```

It became apparent that it is necessary to extend the presented concept for 'multigroup' CHR, an example of which, AC+paclitaxel is defined in NCCN guidelines, see figure 5.

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

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.

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. The structured record of the above illustrated regimen is shown in figure 6 (the dosage of paclitaxel 175mg/m^2 was chosen according to the national guidelines). For multigroup CHR it is necessary that the element *noc* must not be zero, at least for each group except the last.

Figure 6: An example of multigroup CHR definition

```
<regimen>
        <name>ACpac</name>
```

```
<diagnosis>
	<ICD10>C50</ICD10>
	<line>1</line>
	<purpose>adjuvant</purpose>
</diagnosis>
<group>
	<id_group>1</id_group>
	<interval>21</interval>
	<noc>4</noc>
	<drug>
		<name lang="cz">Cyklofosfamid</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>
	<drug>
		<name lang="cz">Doxorubicin</name>
		<name lang="eng">Doxorubicin</name>
		<atc>L01DB01</atc>
		<abbr>A</abbr>
		<adm>
			<day>1</day>
			<dose>60</dose>
			<unit>mg/m2</unit>
			<mode>iv</mode>
		</adm>
	</drug>
</group>
<group>
	<id_group>2</id_group>
	<interval>21</interval>
	<noc>4</noc>
	<drug>
		<name lang="cz">Paklitaxel</name>
		<name lang="eng">Paclitaxel</name>
		<atc>L01CD01</atc>
		<abbr>Pt</abbr>
		<adm>
			<day>1</day>
			<dose>175</dose>
			<unit>mg/m2</unit>
			<mode>iv</mode>
		</adm>
	</drug>
```

```
        </group>
</regimen>
```

For this above described structure an XML scheme was developed using the program XMLSpy version 5.3. Following this scheme individual XML documents were created for each of the selected CHR.

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. 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"

Primary proposal for systematic naming is shown in figure 7.

Figure 7: The structure of CHR systematic name

(Days of administration;Dose;Mode;Unit)Drug_abbreviation +
(Days of administration;Dose;Mode;Unit)Drug_abbreviation 2 +
(Days of administration;Dose;Mode;Unit)Drug_abbreviation 3 &
Interval

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 a Table I.

Abbreviations used are parts of the proposal for the standardization of structures of CHR, because currently there standard definitions for the abbreviated identification of cytostatics have not been found (on the webpages of NCI 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 8.

Figure 8: An example of the identification of the CHR AC+ paclitaxel

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]

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.

Table I - Summary of most frequently used cytostatic agents

Cytostatic agent	ATC - code	NCI abbreviation	Used abbreviation
Bevacizumab	L01XC07	BEVA	Be
Bleomycin	L01DC01	BLEO	B
Busulfan	L01AB01	BU, BUS	Bu
Capecitabine	L01BC06	CAPE	Ca
Carboplatin	L01XA02	CBDCA	Cb
Carmustine	L01AD01	BCNU	Bc
Cetuximab	L01XC06	MOAB C225	Ce
Chlorambucil	L01AA02	CHL, CLB	Cl
Cisplatina	L01XA01	CDDP	P
Cyclophosphamide	L01AA01	CTX	C
Cytarabine	L01BC01	ARA-C	Cy
Dactinomycin	L01DA01	DACT	Ac
Dacarbazine	L01AX04	DTIC	Dc
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
Melphalan	L01AA03	L-PAM	Ml
Methotrexate	L01BA01	MTX	M
Mitomycin C	L01DC03	MITO, MITO-C	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