PhUSE EU Connect 2018

Paper DS03

The oncology specific domains TU, TS and RS: What to know as a statistical analyst

Frieder Wolff, Chrestos Concept GmbH & Co.KG , Essen , Germany
Vitali Gering, Chrestos Concept GmbH & Co.KG , Essen , Germany

ABSTRACT

In oncological clinical trials usually the progression of the disease is of paramount interest. To categorize the cancer response to the medication standardized criteria are applied. Depending on the indication medical experts and statisticians need to choose a suitable approach. As a statistical analyst or data manager it is an advantage to be familiar with the most common criteria. This paper will focus on the RECIST criteria (v1.1) for solid tumors. According to the CDISC Study Data Tabulation Model (SDTM v1.4) the response related data is stored in three tumor domains. To know the coherences between these domains simplifies the derivation of parameters in the analysis data sets that are needed for survival analysis. Furthermore, it helps to identify potential issues early on.

INTRODUCTION

With approximately 14 million new cases in 2012 cancer is one of the leading causes of morbidity and mortality worldwide (Ferlay et al., 2013) and therefore of major interest for the pharmaceutical industry. According to the world preview (EvaluatePharma®, 2017) oncology is the most dominant therapy segment in the pharmaceutical industry with drug & OTC sales of 93.7 $bn and there is no change in course expected.

Clinical trials in oncology have special characteristics, which differ from studies in other therapeutic areas. Typical endpoints as tumor shrinkage or disease progression make it essential to evaluate the treatment success. The demand for standardized criteria for assessing the cancer response in oncological clinical trials arises from a need for comparability, aiding the approval process for new cancer therapeutics by regulatory agencies. Depending on the type of tumor, medical experts and statisticians choose suitable criterion. The first part of this paper shows a brief overview of the most common criteria, in which one is chosen exemplary and therefore described in detail. The second part explains three standardized datasets, which are specific for oncologic trials.

This papers purpose is to establish and explain what to keep in mind as a statistical analyst or data manager working in these domains. The coherences and connections between the individual domains and taking into consideration the medical background are of particular importance.

RESPONSE CRITERIA

The response criteria used in a clinical trial need to be chosen according to the indication analyzed. For example the Revised Response Criteria for Malignant Lymphoma described in (Cheson et al., 2007). The presented guidelines advise on how the response of new cancer therapeutics in non-Hodgkin's and Hodgkin's lymphoma could be assessed, based on positron-emission tomography (PET), immunohistochemistry (IHC) and flow cytometry results. Hallek et al. (2008) updated the guidelines for the diagnosis and treatment of Chronic Lymphociytic Leukemia from the National Cancer Institute. On the basis of blood count, blood smear and the immune phenotype of the circulated lymphoid cells prognostic and diagnostic, the parameters are defined. As RECIST (Eisenhauer et al., 2009) is selected as an example in the following, it will be described in more detail in the next subchapter.

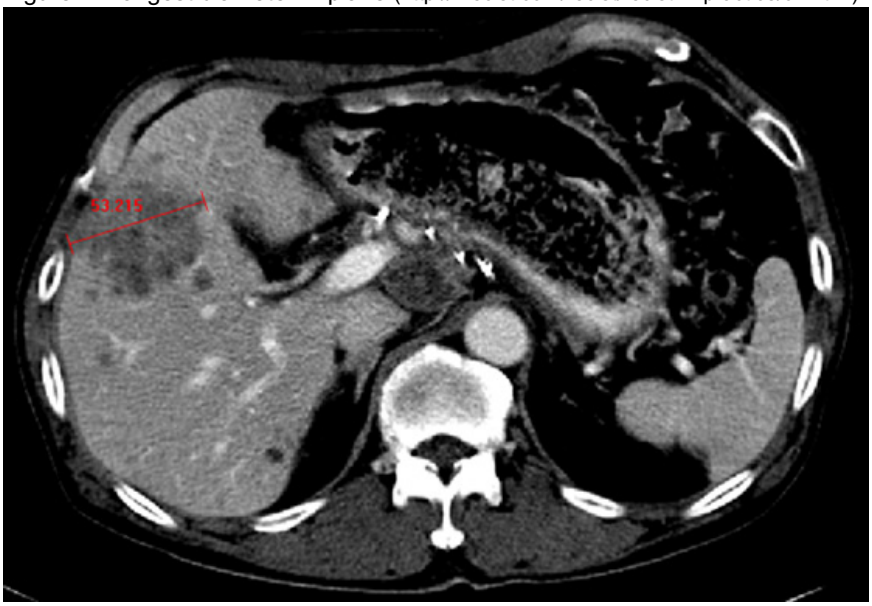
RESPONSE EVALUATION CRITERIA IN SOLID TUMORS (RECIST)

RECIST is a standard method based on anatomical tumor burden, describing an approach to measure solid tumors and assess the change in their size.

It originates from tumor response criteria published by the World Health Organization (WHO) in 1981 and was modified regularly to incorporate new technologies and insights from the ongoing application. In 2000 an International Working Party published new criteria, for the first time named RECIST. Nowadays RECIST is applied in the majority of clinical studies assessing the success of cancer therapies in solid tumors. Such widely accepted criteria are essential to enable comparisons between the endpoints of different studies. Typical tumor endpoints are objective response and time to disease progression.

At the start of a clinical trial a baseline visit is performed with each patient. This visit should be scheduled as close as possible to the start of treatment and not earlier than four weeks before this date to ensure a valid baseline measurement. According to RECIST each tumor lesion or lymph node is categorized *measurable* or *non-measurable* at the baseline visit. Depending on the individual situation CT scan, chest X-ray or caliper measurements by clinical exam could be used to determine the tumor size. The categorized tumor lesions need to be accurately measured in at least one dimension. The measured tumor size is defined as the longest diameter in the plane.

Figure 1: Longest diameter in plane (http://irrecist.com/recist/recist-in-practice/02.html)



Depending on the chosen method the tumor lesion is defined as measurable if the measured size is at least:

- CT scan: 10mm (CT Scan slice thickness no greater than 5mm)
- Caliper measurement: 10mm
- Chest X-ray: 20mm

All lesions that do not fulfill these criteria are considered non-measurable. They are categorized together with the actual non-measurable lesions, such as leptomeningeal disease, ascites, pleural or pericardial effusion, inflammatory breast disease, lymphangitic involvement of skin or lung, abdominal masses/abdominal organomegaly identified by physical exam that is not measurable by reproducible imaging techniques.

As the baseline measurement should be used as reference to assess the tumor response during the trial, the overall tumor burden needs to be estimated. In order to do so the lesions are separated into target and non-target lesions. Up to five measurable lesions per patient, and thereof up to two per organ, could be considered as target lesions. If a patient has more than five measurable lesions, respectively more than two measurable lesions in one organ, the target lesions are to be chosen according to the following description. The target lesions should be those with the longest diameter, and need to be representative for the involved organs and most importantly they should assure reproducible repeated measurements. To ensure those measurements the second or third largest lesion per patient or per organ could be selected as target lesion if the larger lesions appear to be unreliable in this regard. Accurate and therefore comparable measurements throughout the complete trial are absolutely mandatory to assess the response correctly.

Once the target lesions are selected, a sum of diameter is calculated. This baseline sum of diameter will be used as comparator during the study.

The sum is derived from the longest diameter of the non-nodal target lesions and the shortest diameter of the nodal target lesions. As lymph nodes are normal anatomical structures that grow in their width if involved by a tumor, the shortest diameter is taken into account if a pathological lymph node is selected as target lesion.

All lesions and pathological lymph nodes that are detected at baseline but not selected as target lesions are classified as non-target lesions. The measurement of these lesions throughout the trial is not mandatory but their presence or absence should still be reported as these categories are taken into account during the response assessment as well.

The following section describes the different rules to assess the objective tumor response at the visits during the study. Response of tumor is categorized in Complete Response (CR), Partial Response (PR), Stable Disease (SD) or Progressive Disease (PD). For target and non-target lesions different definitions have to be considered. An overall response at each visit will be determined from both results, according to the rules described later. Table 1 contains the definitions of the response categories that are applied in the target lesions.

Table 1: Response evaluation in target lesions

Response	Definition
Complete Response (CR)	Disappearance of all target lesions. Any pathological lymph nodes (whether target or non-target) must have a reduction in short axis to <10 mm.
Partial Response (PR)	At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.
Stable Disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study.
Progressive Disease (PD)	At least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum in the study (this includes the baseline sum if that is the smallest in the study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. (Note: the appearance of one or more new lesions is also considered progression).

As described in the table above the response in target lesions is evaluated based on the sum of diameter. It is important to note that the reference differs between the categories. For PR the decrease of 30 percent needs to be achieved in comparison to the baseline measurement. This reference does not change throughout the study. In contrast to that, the reference for the PD threshold may need to be updated during the trial as the smallest sum in the study is used here. This definition ensures that PD is still detected in target lesions that shrink after the beginning of treatment but begin to grow again afterwards. As long as the sum of diameter does not exceed the thresholds for PR and PD, the response of the target lesions is defined as SD. Independently from the baseline measurement a Complete Response is only reached if all target lesions disappear completely and all pathologic lymph nodes shrink to a normal size.

Definitions of the response categories in the non-target lesions are summarized in Table 2. In comparison to the target lesions the category Stable Disease is replaced by Non-CR/Non-PD and Partial Response is omitted. As mentioned above the non-target lesions do not need to be measured, even if necessary. It is sufficient to assess them categorically.

Table 2: Response evaluation in non-target lesions

Response	Definition
Complete Response (CR)	Disappearance of all non-target lesions and normalization of tumor marker level. All lymph nodes must be non-pathological in size (<10 mm short axis).
Non-CR/Non-PD	Persistence of one or more non-target lesion(s) and/or maintenance of tumor marker level above the normal limits.
Progressive Disease (PD)	Unequivocal progression (described below) of existing non-target lesions. (Note: the appearance of one or more new lesions is also considered progression).

An ‘unequivocal progression’ in the non-target lesions is achieved if the overall level of tumor burden in these lesions increase sufficiently to enforce the discontinuation of therapy. Such a progression could be determined even if the Response in the target lesions is positive.

When the response in target and non-target lesions is assessed according to the rules described above, an Overall Response per tumor assessment visit could be determined. Rules on how both results are to be combined are shown in Table 3. A patient is not evaluable (NE) at a time point if no imaging or measurement has been done.

RECIST describes additional rules in case a patient has only non-target lesions, but this is not covered in this paper.

Table 3: Overall response per time point

Target lesions	Non-target lesions	New lesions	Overall Response
CR	CR	No	CR
CR	Non-CR/Non-PD	No	PR
CR	Not evaluated	No	PR
PR	Non-PD or not all evaluated	No	PR
SD	Non-PD or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

Once all response data per time point is collected and evaluated, the best overall response for a patient can be determined. It is defined as the best overall response that a patient has reached during the trial. Often Complete Responses need to be confirmed by an additional assessment before they are accepted. The specific rules for a confirmation are to be defined in the study protocol, but RECIST advises how this could be handled.

ONCOLOGY SPECIFIC DOMAINS

To capture all relevant information on tumor lesions and the disease response the Submission Data Standards team of Clinical Data Interchange Standard Consortium (CDISC) has developed a standard data structure, which is described in the Study Data Tabulation Model Implementation Guide (SDTMIG) v3.2 (CDISC Submission Data Standards Team, 2013). The respective tumor package includes three SDTM domains: TU, TR and RS. Although they are all related, each domain has its own distinct purpose, serving to represent the data collected in clinical trials (Eben J, 2014). The following chapters describe each of the three domains separately as well as their connection.

TUMOR IDENTIFICATION (TU):

The purpose of the Tumor Identification (TU) domain is to represent data that uniquely identifies tumors, which are tracked during the course of a study. Usually this initial identification is performed by an assessor (e.g. Investigator or Radiologist) at a baseline visit. This identification must not be repeated at every visit, hence only one record per unique tumor is allowed, but there are some examples for including post-baseline records in the TU domain:

- entirely new emerged tumors
- tumors identified at baseline that subsequently split into separate distinct tumors
- multiple tumors that merge together, leading to the inclusion of an additional post-baseline record to distinctly identify the merged tumor

The following section describes the most important variables used in the TU domain. The variable TULNKID is used as a link between the identified tumors (TU domain) and the assessment results over the course of the study (TR domain), using a unique code for each identified tumor. The test for the tumor identification is represented using TUTEST and TUTESTCD (short name of the test in TUTEST). In case of a tumor identification the value “TUMIDENT” gets assigned to TUTESTCD. If a tumor happens to split into one or more distinct tumors or two or more tumors merge to form a single tumor, this information is stored in the TU domain as well. A block of related records (e.g. split or merged tumors) within a subject is linked through a group ID (TUGRPID). Such new records are created with TUTEST set to ”Tumor Split” or “Tumor Merge”. The anatomical location of the identified tumor is specified in the variable TULOC whereas the used method of assessment (e.g. MRI, CT) is collected in the variable TUMETHOD. The result of the tumor identification is stored in TUORRES and represents a classification of the identified tumor. When TUTESTCD equals ”TUMIDENT” exemplary values of TUORRES are TARGET, NON-TARGET or NEW. The role of the evaluator the lesions (e.g. INVESTIGATOR or INDEPENDENT ASSESSOR) is collected in TUEVAL. Table 4 shows an example of how the above described variables could be filled in for one subject.

Table 4: Important variables used in the TU domain filled in exemplarily for one subject

ROW	USUBJID	TULNKID	TUGRPID	TUTESTCD	TUTEST
1	44444	T01		TUMIDENT	Tumor Identification
2	44444	T02		TUMIDENT	Tumor Identification
3	44444	T03		TUMIDENT	Tumor Identification
4	44444	T04		TUMIDENT	Tumor Identification
5	44444	NT01		TUMIDENT	Tumor Identification
6	44444	NT02		TUMIDENT	Tumor Identification
7	44444	T04.1	T04	TUSPLIT	Tumor Split
8	44444	T04.2	T04	TUSPLIT	Tumor Split
9	44444	T02/T03		TUMERGE	Tumor Merged
10	44444	NEW01		TUMIDENT	Tumor Identification

ROW (cont.)	TUORRES	TULOC	TUMETHOD	TUEVAL	VISIT
1	TARGET	LIVER	CT SCAN	INVESTIGATOR	SCREEN
2	TARGET	KIDNEY	CT SCAN	INVESTIGATOR	SCREEN
3	TARGET	KIDNEY	CT SCAN	INVESTIGATOR	SCREEN
4	TARGET	LUNG	CT SCAN	INVESTIGATOR	SCREEN
5	NON-TARGET	THYROID GLAND	CT SCAN	INVESTIGATOR	SCREEN
6	NON-TARGET	CEREBELLUM	MRI	INVESTIGATOR	SCREEN
7	TARGET	LUNG	CT SCAN	INVESTIGATOR	WEEK 6
8	TARGET	LUNG	CT SCAN	INVESTIGATOR	WEEK 6
9	TARGET	KIDNEY	CT SCAN	INVESTIGATOR	WEEK 24
10	NEW	CERVICAL LYMPH NODE	MRI	INDEPENDENT ASSESSOR	WEEK 32

To indicate that only one subject is considered for the above example, a unique subject ID (USUBJID) containing the same value (“44444”) for all rows is displayed. Not only that TULNKID functions as a linking variable between the domains, its value also contains information regarding the lesions classification. The first tumor classified as target gets assigned to “T01” whereas the first non-target lesion is represented by “NT01”. Furthermore, the variable VISIT contains information about when the finding was documented. The first six entries in the above example are tumor identifications, hence the respective VISIT is the screening. On week 16 a tumor split of the lesion “T04” was determined, leading to two additional records having link IDs “T04.1” and “T04.2” respectively, while the “TUGRPID=T04” shows the connection to the original identified lesion.

TUMOR RESULTS (TR):

The Tumor Results (TR) domain contains information about tumors identified in the TU domain in form of quantitative measurements and/or qualitative assessments of the tumors. These measurements are not only taken at baseline but at each subsequent assessment with the purpose of supporting the response evaluations. In contrast to the TU domain the TR domain does not contain information about anatomical locations of the measurements.

Some variables like TRLNKID, TRGRPID, TREVAL or TRMETHOD function in the same way as their respective counterparts from the TU domain. For the sake of consistency TRMETHOD and TUMETHOD are (in most cases) the same throughout the trial. Using the variable TRLNKGRP enables the grouping and linking of all assessment records used in the assessment of the response record in the RS domain. Information regarding the test that has been used to obtain the measurement or finding (e.g. tumor state, diameter) is stored in TRTEST/TRTESTCD. It is of importance that the sponsor should not derive results for any test if the result was not collected, i.e. the inclusion of tests in this domain would only be allowed if the respective data has been collected on a CRF or provided by an external assessor as part of an electronic data transfer. The result of the tumor measurement in its originally obtained form is stored in TRORRES while TRORRESU contains the corresponding unit. Whenever an assessment was not performed or a tumor measurement was not taken, the variable TRSTAT is used as an indicator. As a supplement TRREASND describes the reason for the undone assessment. Both of these variables should be Null if a result exists in TRORRES.

Table 5 includes the subject`s findings from the previously used example up to week 12 and shows how a possible continuation in TR could be realized.

Table 5: Important variables used in the TR domain, proceeding the captured data of the exemplarily used subject (only up to week 12). The respective entries for TRTEST would be Diameter (DIAMETER), Tumor State (TUMSTATE), Sum of Diameter (SUMDIAM), Percent Change from Baseline in Sum of Diameter (PCBSD) and Percent Change from Nadir in Sum of Diameter (PCNSD).

TRLNKID	TRLNKGRP	TRTESTCD	TRORRES	TRORRESU	TRSTAT	TRREASND	VISIT
T01	A1	DIAMETER	17	mm			SCREEN
T02	A1	DIAMETER	12	mm			SCREEN
T03	A1	DIAMETER	18	mm			SCREEN
T04	A1	DIAMETER	14	mm			SCREEN
	A1	SUMDIAM	61	mm			SCREEN
NT01	A1	TUMSTATE	PRESENT				SCREEN
NT02	A1	TUMSTATE	PRESENT				SCREEN
T01	A2	DIAMETER	0	mm			WEEK 6
T02	A2	DIAMETER	10	mm			WEEK 6
T03	A2	DIAMETER	15	mm			WEEK 6
T04	A2	DIAMETER			NOT DONE	TUMOR SPLIT OR DIVIDED	WEEK 6
T04.1	A2	DIAMETER	6	mm			WEEK 6
T04.2	A2	DIAMETER	7	mm			WEEK 6
	A2	SUMDIAM	38	mm			WEEK 6
	A2	PCBSD	-38	%			WEEK 6
	A2	PCNSD	-38	%			WEEK 6
NT01	A2	TUMSTATE	ABSENT				WEEK 6
NT02	A2	TUMSTATE	ABSENT				WEEK 6
T01	A3	DIAMETER	0	mm			WEEK 12
T02	A3	DIAMETER	15	mm			WEEK 12
T03	A3	DIAMETER	14	mm			WEEK 12
T04	A3	DIAMETER			NOT DONE	TUMOR SPLIT OR DIVIDED	WEEK 12
T04.1	A3	DIAMETER	7	mm			WEEK 12
T04.2	A3	DIAMETER	6	mm			WEEK 12
	A3	SUMDIAM	42	mm			WEEK 12
	A3	PCBSD	-31	%			WEEK 12
	A3	PCNSD	11	%			WEEK 12
NT01	A3	TUMSTATE	ABSENT				WEEK 12
NT02	A3	TUMSTATE	ABSENT				WEEK 12

As the sum of diameter falls from 61 mm at baseline to 38 mm at week 6, the percent change results in a decrease of 38 %. Since the screening is the only visit prior to week 6, it is also used as reference for calculating the percent change from nadir, resulting in a decrease of 38 % as well. The percent change from baseline to week 12 in sum of diameter equals to -31 % due to an increase in sum of diameter from 38 mm to 42 mm comparing week 6 and 12 respectively. For the calculation of the percent change from nadir in sum of diameter at week 12 the smallest value up to this point, which is the sum of diameter at week 6 (38 mm), is chosen. This leads to an increase of 11 %.

DISEASE RESPONSE (RS):

The response evaluation determined based on the information from TR is collected in the RS domain. Additional input (e.g. lab test results) may also be considered for assessing the response as all available relevant information is used depending on the response criteria.

To enable a link between the response assessment and the records from TR which were used in the assessment of the response, RSLNKGRP is to be used. The type of response assessment (e.g. overall response, response of measurable lesions) is represented by RSTEST/RSTESTCD while RSORRES contains the result of the response assessment that originally has been received. As previously described in TR, RSSTAT is used to indicate if a response assessment was not performed whereas RSREASND describes the reason why the assessment was not performed. The criteria used in the assessment of response including a version number is captured in RSCAT.

Table 6: Important variables used in the RS domain, proceeding the captured data of the exemplarily used subject (only up to week 12)

RSLNKGRP	RSTEST	RSCAT	RSORRES	VISIT
	Target Response	RECIST 1.1	PR	WEEK 6
	Non-Target Response	RECIST 1.1	CR	WEEK 6
A2	Overall Response	RECIST 1.1	PR	WEEK 6
	Target Response	RECIST 1.1	PR	WEEK 12
	Non-Target Response	RECIST 1.1	CR	WEEK 12
A3	Overall Response	RECIST 1.1	PR	WEEK 12

The target response at week 6 is PR, because the decrease in the sum of diameters (taking the baseline sum diameters as reference) is 38% and therefore greater than 30%. Although the percent change from nadir equals 11% at week 12, the target response is still PR. This is due to the decrease in percent change from baseline, which is 31%, hence still greater than 30%. The non-target response at both week 6 and 12 is CR, as the non-target lesions are absent after the initial finding at baseline. According to table 3 the overall response is PR at both time points.

CONCLUSION

Working as a statistical programmer in the therapeutic area of oncology, it is of importance to have an idea of the purpose and the structure of each of the three tumor domains. To comprehend how the connection between the different domains and the derivation of variables that contribute to determine overall responses based on standardized rules works, helps in developing an understanding of the bigger picture. These rules differ depending on the response criterion used, and this paper is not intended to cover all possible constellations. Instead it gives an idea on how a real-life scenario could look like by using the RECIST 1.1 criterion.

REFERENCES

Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, et al. Revised Response Criteria for Malignant Lymphoma. Journal of Clinical Oncology. Vol 25 Number 5. 2007.

Clinical Data Interchange Consortium (CDISC) Submission Data Standards Team. Study Data Tabulation Model Implementation Guide: Human Clinical Trials. Version 3.2. 2013

Eben J. Implementation of oncology specific SDTM domains. Paper PP06. Phuse 2014.

Eisenhauer E.A., Therasse, P., Bogaerts, J., Schwartz, L.H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D., Verwij, J. New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1). European Journal of Cancer. Vol 45. 2009; (http://www.eortc.be/recist/)

EvaluatePharma®. Word Preview 2017, Outlook to 2022. 10th Edition June 2017. (http://info.evaluategroup.com/rs/607-YGS-364/images/WP17.pdf)

Ferlay J, SoerjomataSTram I, Ervik M, Dikshit R, Eser S, Mathers C et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11, Lyon, France: International Agency for Research on Cancer; 2013

Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H, Hillmen P, Keating MJ, Montserrat E, Rai KR, Kipps TJ. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia (IWCLL) updating the National Cancer Institute-Working Group (NCI-WG) 1996 guidelines. Blood. Vol 111. 2008.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Frieder Wolff
Chrestos Concept GmbH & Co.KG
Girardetstr. 1-5
45131 Essen
Work Phone: +49 175 / 2139780
Email: frieder.wolff@chrestos.de
Web: http://www.chrestos.de

Vitali Gering
Chrestos Concept GmbH & Co.KG
Girardetstr. 1-5
45131 Essen
Work Phone: +49 160 / 97761157
Email: vitali.gering@chrestos.de
Web: http://www.chrestos.de