eDESDE-LTC

DESCRIPTION AND EVALUATION OF SERVICES AND DIRECTORIES IN EUROPE FOR LONG TERM CARE

USABILITY OF THE eDESDE-LTC INSTRUMENT:

Feasibility, Consistency, Reliability and Validity

Luis Salvador-Carulla, Juan Luis González-Caballero, Mencía Ruiz, Daniela Seyrlehner, María Teresa Roma-Ferri, Miriam Poole, Carolina Lagares, Ana Fernandez, Cristina Romero, *for the eDESDE-LTC Group*

Executive Agency for Health and Consumers (EAHC)

Project Ref. 2007/116

Executive Agency for Health and Consumers

eDESDE-LTC Project: Word Package WP4 – Deliverables D13 and D15

Usability of the eDESDE-LTC Instrument: Feasibility, Consistency, Reliability and Validity

Luis Salvador-Carulla

Juan Luis Gonzalez-Caballero

Mencía Ruiz

Daniela Seyrlehner

Maria Teresa Roma-Ferri

Miriam Poole

Carolina Lagares

Ana Fernandez

Cristina Romero

for the eDESDE-LTC Group

This document should be quoted as follows:

Salvador-Carulla L, Gonzalez-Caballero JL, Ruiz M, et al. for the eDESDE-LTC Group. eDESDE-LTC: Usability of eDESDE-LTC [Internet]. Jerez (Spain): PSICOST and Telnet; 2011. Available from: http://www.edesdeproject.eu

---

Edition: January 2011

Edits: PSICOST

Plaza San Marcos nº 6, 11403, Jerez de la Frontera (Spain)

Electronic edition and design: Telnet Consulting S.L.

ISBN: (requested)

eDESDE-LTC Partners

Asociación Científica PSICOST (Spain), London School of Economics (LSE) (UK), SINTEF Technology and Society (Norway), Catalunya Caixa (Spain), Družbenomedicinski inštitut ZRC SAZU (Slovenia), Public Health Association (PHA) (Bulgaria), University of Vienna (UNIVIE) (Austria), Izobraževalno raziskovalni inštitut Ljubljana (IRIO) (Slovenia)

Collaborating Partners

Dept. of Mental Health Sciences, UCL (UK); ASL-4 Centro Studi e Ricerche (Italy), Dept Research & Development, Division of Mental Health Serv. Akershus University Hospital (Norway), ETEA University of Cordoba (Spain), University of Alacant (Spain), University Politecnica of Barcelona (Spain), OECD – Health Division (France), Ministry of Health. National Center for Public Health Protection (Bulgaria), Catalan Department of Health, DG de Planificació i Avaluació (Spain), Psychiatry Research Unit of Cantabria (WHO collaborating Centre) Univ. Hosp."Marqués de Valdecilla" (Spain), Fundació Villablanca (Spain), Delegación Municipal de Igualdad y Salud. Ayto Jerez (Equity & Health Department. Municipality) (Spain).

Acknowledgements

The eDESDE-LTC project has been funded by the Executive Agency of Health and Consumer (EAHC) Project Ref. 2007/116

eDESDE-LTC Group *(members[1], collaborating members[2] and main subcontractors[3])*

José Almenara (Spain)[1]
Federico Alonso (Spain)[1]
Jordi Balot (Spain)[1]
Murielle Bendeck (Spain)[3]
Anne Mette Bjerkan (Norway)[1]
Barbara Brehmer (Austria)[1]
Angel Broshtilov (Bulgaria)[1]
Maria Dolores Crespo (Spain)[2]
Mojca Z. Dernovsek (Slovenia)[1]
Hristo Dimitrov (Bulgaria)[1]
Nikolina Djurova (Bulgaria)[1]
Mónica Duaigües (Spain)[1]
Josep Fuste (Spain)[2]
Ana Fernandez (Spain)[3]
Luis Gaite (Spain)[2]
Carlos García-Alonso (Spain)[2]
Juan Carlos García Gutierrez (Spain)[1]
José García-Ibañez (Spain)[2]
Juan Luis Gonzalez-Caballero (Spain)[1]
Karina Gibert (Spain)[2]
Sonia Johnson (UK)[2]
Birgitte Kalseth (Norway)[1]
Martin Knapp (UK)[1]
Carolina Lagares (Spain)[1]
Paula Llull (Spain)[1]
Teresa Marfull (Spain)[3]
Rafael Martinez-Leal (Spain)[1]
Tihana Matosevic (UK)[1]
David McDaid (UK)[1]
Cristina Molina (Spain)[2]
Valerie Moran (France)[2]
Carlos Mur (Spain)[2]
Carmen Omist (Spain)[2]
Miriam Poole (Spain)[3]
Manuel Palomar (Spain)[2]
Rayna Popova (Bulgaria)[1]
Eugeni Rojas (Spain)[1]
Maria Teresa Romá-Ferri (Spain)[2]
Cristina Romero (Spain)[1]
Mencía Ruiz (Spain)[1]
Torleiff Ruud (Norway)[2]
Jose Alberto Salinas (Spain)[1]
Luis Salvador-Carulla (Spain)[1]
Antoni Serrano (Spain)[3]
Daniela Seyrlehner (Austria)[1]
Josep Solans (Spain)[1]
Lilijana Sprah (Slovenia)[1]
Giuseppe Tibaldi (Italy)[2]
Jose Luis Vazquez-Barquero (Spain)[2]
Britt Venner (Norway)[1]
Germain Weber (Austria)[1]
Elisabeth Zeilinger (Austria)[1]

CONTENTS

FOREWORD

The 'Description and Evaluation of Services and Directories in Europe for Long Term Care' (DESDE-LTC) is an instrument for the standardised description and classification of services for Long-Term Care (LTC) in Europe. DESDE-LTC has been designed to allow national and international comparisons.

The eDESDE-LTC usability report provides information on the psychometric properties of the instrument, including feasibility, consistency, reliability and validity. This report is available at http://www.edesdeproject.eu[1].

Luis Salvador-Carulla

Coordinator of eDESDE-LTC Project

[1] If you want to provide us a feedback on the usability of the eDESDE-LTC system, please click on the link below to complete the online questionnaire (it takes less than 10 minutes):

http://www.unet.univie.ac.at/~a0305075/umfragen/index.php?sid=21575&newtest=Y&lang=en

LIST OF MAIN ABBREVIATIONS

BSIC	Basic Stable Inputs of Care
DESDE	Description and Evaluation of Services and Directories
EAHC	Executive Agency of Health and Consumers
EASPD	European Association of Service Providers for Persons with Disabilities
ELSA	English Longitudinal Study of Ageing
ESMS	European Service Mapping Schedule
EQM	Evaluation Quality Management
EPCAT	European care Psychiatric Assessment Team
IHTSDO	International Health Terminology Standards Development Organisation
IRIO	Izobraževalno Raziskovalni Inštitut
LSE	London School of Economics
LTC	Long-Term Care
MHEEN	Mental Health Economic European Network
MTC	Main Types of Care
OECD	Organisation for Economic Co-operation and Development
QAP	Quality Assessment Plan
SEP	Sociedad Española de Psiquiatría
SHA	Public Health Association
UNIVIE	University of Vienna
WHO	World Health Association

LIST OF TABLES AND BOXES

TABLES

1. INTRODUCTION

Evaluation and assessment are essential components of healthcare, and they require assessment instruments with known metric properties. However, the foundation of health metrics has been developed in a scattered way and the related knowledge is still fragmented, with uneven development in different areas. The evaluation of intangible phenomena (pain, anxiety, disability, quality of life, quality of care), raised a whole array of complex questions with regard to feasibility, consistency, validity and cultural transferability, among others. In any case, a considerable effort towards harmonisation has been produced, particularly in item analysis and standardisation of instruments and quality measures such as validity, reliability, feasibility, usability, and comparability, as well as links between care provision and outcomes (Ishak et al, 2002; Furr & Bacharach, 2008; Salvador-Carulla & Gonzalez-Caballero, 2010).

However the assessment of the psychometric properties of instruments designed for health service research and planning has received less attention than those instrument aimed at assessing patient's status, functioning, satisfaction or preferences. This can be partly attributed to the variability of domains and related instruments in this area (Lloyd-Evans et al, 2007; IHFAN, 2008). Some information is available on care utilisation instruments such as the Client Service Receipt Inventory (CSRI) (Chisholm et al, 2000), or the Resident Assessment Instrument-Mental Health (RAI-MH) (inter-rater reliability and convergent validity) (Hirdes et al, 2002). Previous studies on the psychometric properties of instruments for assessment of availability and use of services for territorial comparisons include the European Service Mapping Schedule (ESMS) for the assessment of mental health services (Salvador-Carulla et al, 2000) and its adaptation for assessment of services for persons with disabilities in Spain (Salvador-Carulla et al, 2006).

The previous literature on the quality parameters of instruments for health service availability is scarce. Therefore a review of the relevant aspects is here provided.

Feasibility

Feasibility has become a relevant issue in health assessment particularly in health care (Salvador-Carulla & Gonzalez-Caballero, 2010). There is no consensus on how feasibility should be defined and measured. Slade and co-workers have suggested a definition in the context of routine outcome assessment (Slade et al, 1999). Andrews et al. (1994) identified three dimensions of feasibility: applicability, acceptability, and practicality.

The *applicability* of a measure was defined as the degree to which a measure addresses dimensions of importance to the consumer, is useful for services providers in formulating and conducting decisions, an allows for the aggregation of data in a meaningful way to meet the purposes of service management. This aspect, defined as “relevance” by Slade and colleagues (1999), may be framed as follows: Is the description meaningful to recipients? (ex., to health authorities, managers, staff, patients/families). The *acceptability* of a measure describes the ease with which a consumer or clinician can use a particular measure (i.e., user-friendliness). *Practicality* relates to implementation, training requirements, and complexity of scoring, reporting and interpreting the data. Efficiency may be regarded as the fourth dimension of feasibility. It could be defined as the relationship existing between its practicality and the costs incurred by its utilization.

Consistency (structural validity or internal reliability)

Consistency comprises the psychometric solidity of a scale, its internal structure, the level to which its different items are interrelated and the possibility of adding them up to obtain overall scores. It has been defined as that property which defines the level of agreement or conformity of a set of measurements among themselves. Some authors include consistency within the category of reliability and it is also related to its structural validity.

However, the internal reliability or consistency should be clearly differentiated from the external reliability of an assessment instrument.

Homogeneity, indicates the degree of agreement among the items on a scale, which determines if they can accumulate and generate an overall score. It can be obtained by studying the correlation of the items with the total (using split-half reliability, Kuder Richardson formula KR-20 or Cronbach’s alpha), by factor analysis or by using Rasch’s statistical objectivity models. Homogeneity based on factor analysis (acceptability of the global score as the sum of that obtained on each item) is confirmed if a one-dimensional structure is obtained, that is, all the items show a positive load on the first factor. In addition to exploratory factor techniques such as principal component analysis and principal factor analysis, the structure of a scale can be assessed using other techniques, such as the non-metric multidimensional scaling or the structural equation analysis (Salvador-Carulla & Gonzalez-Caballero, 2010).

Formal ontology provides a new perspective on the content structure of assessment instruments, particularly in the field of classification systems. Formal ontology is an explicit specification of a conceptualization (the objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them) (Gruber, 1993). It has been applied to computer sciences to formalize the concepts, hierarchies and the existing relationships among different concepts in order to enable semantic interoperability and data transfer. These techniques provide a sound method to describe the content architecture or hierarchy of any classification and assessment instrument, identifying internal errors and thus describing the content consistency of the instrument (for example, Heja et al, 2008). An ontology approach has been previously applied to the analysis of home care for the elderly in Spain (Valls et al, 2010). A formal ontology analysis is essential to facilitate the semantic interoperability of any classification and/or coding system (Roma-Ferri et al, 2005).

Reliability (external reliability)

"Reliability" reflects the amount of error, both random and systematic, inherent in any measurement procedure. It has been defined as the proportion of variance in a measurement that is not error variance, excluding errors related to consistency (attributable to the internal structure of the instrument). In this sense, the reliability will provide information about the reproducibility of the test's results in different situations, or also, it will indicate the degree of the stability of the test's measures, in spite of changes in different external parameters (that is, not inherent in the test). There is a wide terminological variability in the terms and methods used to assess the reproducibility or the stability of assessment instruments (i.e. accuracy, precision, agreement, dependability and consistency) (Salvador-Carulla & Gonzalez-Caballero, 2010). Reliability could be framed in the context of the Classical Test Theory (CCT) or in the context of the Generalisability Theory (GT). The different approaches and their related statistical techniques have been previously reviewed (Salvador-Carulla & Gonzalez-Caballero, 2010). It should be noted that GT allows the simultaneous analysis of several coefficients of reliability (inter- and intra-observer, test-retest, inter-informant, etc) which could be generalised to fixed or to random conditions. As an example, a questionnaire of service utilisation may be used by different observers to collect data from different sources (clinical records, patients, family carers). Using GT we may be able to assess the reliability of every section and the overall questionnaire in different groupings of observers and information sources, selecting the combination of facets which provide the higher reliability.

Validity

Validity indicates which proportion of the information collected is relevant to the formulated question, and is defined by the degree to which an instrument measures what it is supposed to measure. Validity and reliability are closely connected. On the one hand, validity cannot be assessed unless the instrument is reliable. On the other, reliability and validity are related in the decomposition of the observed variance of scale's scores. It includes the random error, the construct variance, and the variance due to systematic errors (Judd et al, 1991). The construct variance has a direct influence on the validity and the reliability of any instrument, whereas systematic errors influence only the reliability. Optimizing both reliability and validity requires sacrificing the maximization of each (attenuation paradox).

Validity is considered present when the measurement predicts a criteria (criterion validity), or consistently fits a series of related constructs within the context of an accepted theory (construct validity), if there is no external criterion that serves as a gold standard. There are multiple forms of validity, with the further complication that some authors use the same term to define different concepts. The six main forms of validity can be distributed into two axes: one revolving around the presence or absence of a gold standard for the dimension assessed (criterion validity vs construct validity), and another focus on whether mathematical techniques are used in their calculation (descriptive validity vs statistical validity). Thus, a certain type of validity can be considered of the criterion or the construct type, depending on the dimension assessed. Concurrent validity of a scale for services assessment forms part of criterion validity, whereas the concurrent validity for a quality-of-life scale, for which there is no gold standard, should be considered as part of its construct validity. Likewise, estimation of discriminant validity or convergent validity may be merely descriptive, or may involve use of statistical procedures. The principle types of validity of an assessment instrument have been previously reviewed (Salvador-Carulla and Gonzalez-Caballero, 2010):

Simple validity (face validity). This is a type of descriptive criterion validity, which reflects what experts consider significant measures. There may be a certain amount of confusion between this concept and that of applicability and relevance (regarded as feasibility domains). It could be mentioned that the latter refers to the judgment of a wide-ranging group of users of the instrument, or of the information derived from it (e.g. healthcare managers or clinicians), whereas the assessment of face validity is limited to the expert´s opinion.

Content validity. Defines the degree to which the set of items on a test adequately represents the domain assessed, i.e., the level of representativeness of the items of the set of components under assessment. In reality, this concept does not differ much from that of consistency, so that they may be considered synonymous. According to Thompson (1989), this type of validity is also descriptive, and cannot be analyzed using statistical techniques. Formal ontology has provided a new perspective to the analysis of content validity, particularly in the field of classification systems (Romá-Ferri & Palomar, 2008).

Commensurability is other concept closely linked to content validity. Commensurability is related to the “apples and oranges” problem, where substantively disparate items have been grouped together. Classical measures of quality of life may face content validity problems related to this factor (Steel et al, 2008)

Discriminant validity. This refers to the degree to which an instrument measures those features belonging to one domain and not to others, as well as the degree to which the features of different domains are not included within the domain examined by the instrument (inclusion and exclusion discriminant validity). Discriminant validity may be assessed either descriptively or with statistical procedures.

Convergent validity. This refers to the assessment of a certain feature of a domain with two different methods (e.g. assessment of depression using an assessment scale and a biological test). This term has also been used to denote the use of two assessment instruments, each covering a different dimension, in order to find a third (e.g. use of clinical and functioning scales to study the validity of a quality-of-life scale).

Concurrent validity. This provides a measure of the association between the scores for different items and the overall scores for other reference scales with an equivalent purpose and content. It is generally limited to the study of inter-score correlation.

Predictive validity. Predictive observation validity refers to the probability that a scale gives a correct judgement of the observed phenomenon. The use of Bayes's analysis makes it possible to determine the predictive validity of a test, its utility and its comparability, based on an analysis of the distribution of 'cases' and 'non-cases' in a given population, as well as its relationship with the results obtained on the test under study (positive or negative). In this case predictive validity is not applicable..

This study is aimed at describing the usability of the eDESDE-LTC system (instrument and coding system) thorough the analysis of the following quality domains: feasibility, consistency, reliability and validity.

2. METHODS

Once the final eDESDE-LTC versions of the instrument and the coding system were available, the usability of the eDESDE-LTC system was analyzed according to four quality parameters: Feasibility, Consistency, Reliability and Validity (Salvador-Carulla & Gonzalez-Caballero, 2010). As the instrument uses the coding system, consistency, validity and reliability provide information on both parts of the eDESDE-LTC system whilst feasibility mainly refers to the instrument.

The feasibility analysis was carried out by the University of Vienna. The full description of the feasibility is available at the eDESDE-LTC Evaluation and Quality Assessment report (Zeilinger et al, 2011). A summary of the feasibility of the instrument is provided here.

The consistency and validity analyses were carried out by the PSICOST research association with the University of Cadiz (Spain). The Sant Joan de Deu Foundation (Spain) contributed to the reliability analysis, and the University of Alacant (Spain) to the qualitative consistency analysis.

Sample

The analysis was made on the PSICOST database of mental health services. This database includes full information on services from different regions in Spain. It is not limited to health services and it includes also services from other sectors related to mental health care such as social, education, work and crime and justice services. A series of services for other LTC groups were purportedly selected for the reliability and validity exercises in order to cover a range of MTC as broad as possible. This set was selected from the PSICOST database on services for disabilities. The case vignettes based on actual services and provided by other 5 European countries were also included.

Feasibility analysis

An *ad-hoc* instrument was designed by the University of Vienna group to assess the feasibility of eDESDE-LTC (Seyrlehner, 2010). The feasibility questionnaire followed the approach developed by Andrews (1994) and Slade et. al (1999). This feasibility evaluation tool included four domains: Applicability, Acceptability, Practicality and Relevance. The latter is closely related to face validity (Salvador-Carulla and Gonzalez-Caballero, 2010). It is not only seen as the construct fulfilling the criteria of feasibility best, but is also considered by survey participants as the most important construct for the assessment of the feasibility of DESDE- LTC.

For creating this 23-items questionnaire a 5-point likert scale was used (1=best/highest / 5= worst/lowest judgment). The participants had also the possibility to give further comments to each question, or giving the answer “the question is unclear to me” or “no answer”. This questionnaire was available on-line.[2] It was completed by members of the partner groups and nominal group participants as well as by health service researchers with previous experience in the use of ESMS/DESDE.

An analysis of the consistency and usability of the feasibility evaluation questionnaire was carried out in the preliminary sample (21 respondents). It showed good internal consistency (Cronbach’s alpha over 0,7 in all domains) (Cronbach et al., 1972). Only three questions out of 23 raised some problems of understanding. The questionnaire covered main aspects of feasibility according to all the experts’ opinion (high content validity) (Seyrlehner, 2010).

Consistency

The qualitative analysis of the hierarchy of the eDESDE-LTC coding system was made by an ontology expert (MR-F) based on previous experience in the ontology analysis of other health classification systems (Roma-Feri and Palomar, 2008). The full explanation of the procedure is described elsewhere (Roma-Ferri, 2009). The relationships of the different terms on the hierarchy was appraised according to their attributes was analyzed here are many ways in which terms can relate to each other in a hierarchy, depending on the attributes of the concept of interest: 1) structural assemble: ‘part-of’ (part-whole), similarity: ‘is-a’ (kind-of, or causal (to explain) how a chain of events could unfold. These types of links allow one term to inherit properties from other terms higher up in the hierarchy. What is inherited depends entirely on the type of link.

- In a ‘part-of’ hierarchy, terms inherit their location from parent terms higher in the hierarchical tree.
- In a ‘is-a’ (kind-of) hierarchy many different properties of parent terms are inherited by their children terms.

[2] Available at the eDESDE-LTC website (http://www.edesdeproject.eu) and at University of Viena: http://www.unet.univie.ac.at/~a0305075/umfragen/index.php?sid=21575&newtest=Y&lang=en

A proxy quantitative analysis of the overall consistency of the instrument was obtained by assessing the association of codes, stability and independence across the three levels at the Boolean factorial analysis (see below).

Reliability analysis

To carry out the reliability analysis, 170 services covering main types of care in Europe were selected by one member of the group (MP) from the Spanish eDESDE database and the case vignettes provided by other European partners. This list included services for mental health, intellectual disabilities, physical disabilities and elderly population. All services were coded according to DESDE-LTC branches (I, A, S, O, D, R; that is Information, Accessibility, Self-support, Outpatient, Day, and Residential care) by two judges Alpha and Beta, where Alfa represents an experienced person on the use of the instrument and Beta a non experienced person.

The reliability analyses took into account both the Classical Test Theory and the Generalizability theory (G theory) (Salvador-Carulla and Gonzalez-Caballero, 2010). The focus of classical test theory (CTT) is on determining error of the measurement but it only allows to estimate one type of error at a time. Essentially it throws all sources of error into one error term. This may be suitable in the context of highly controlled laboratory conditions, but in field research, it is unrealistic to expect that the conditions of measurement will remain constant. The Cohen's Kappa coefficient has been used to provide a measure of the degree to which two judges, A (Apha) and B (Beta), concur in their respective sortings of *n* items into *k* mutually exclusive categories.

Generalizability theory (G Theory) acknowledges and allows for variability in assessment conditions that may affect measurements. The advantage of G theory lies in the fact that it is possible to estimate what proportion of the total variance in the results is due to the individual factors that often vary in assessment, such as setting, time, items, and raters. Another important difference between CTT and G theory is that the latter approach takes into account how the consistency of outcomes may change if a measure is used to make absolute versus relative decisions. In G a universe, its facets, and the conditions for admissible observations are defined through careful construct explication, the traditional domain of validity theory. Given a particular universe of admissible observations, a person's universe score (μ_p) can be defined as the average score based on all admissible observations (X_p) of the universe of interest. The purpose of a measurement is to accurately estimate this universe score ($\hat{\mu}_p$) based on a sample of observations.

Validity analysis

It should be mentioned that the feasibility analysis includes several items that may be regarded also as descriptive/criterion validity domains. To avoid redundancy this domains related to content and to face validity have been analyzed within the feasibility analysis (Zeilinger et al 2011). At the feasibility questionnaire items related to face validity are Section B, Applicability, Question B1: ‘In your opinion, is the data obtained when applying the instrument useful?’ and Section E: ‘Relevance’. Another item is related to content validity: Section B, Applicability, Question B3:’From your point of view, does the instrument cover important dimensions?’.

The quantitative validity analysis of the eDESDE-LTC instrument was made on a database comprising 1339 services. This included services from different regions in Spain (mostly on mental health care) as well as the case vignettes based on real settings by other European countries. This sample covered services for mental health, intellectual disabilities, physical disabilities and elderly population.

Boolean factor analysis was used to evaluate the content validity (the degree to which the set of items on a test adequately represents the domain assessed, i.e., the level of representativeness of the items of the set of components under assessment) and the concurrent validity (the degree to which results from one test agree with results from other, different tests) of DESDE-LTC instrument.

This analysis differs from that of classical factor analysis on binary valued data even though the goal and model (symbolically) appear similar. The goal is to express *p* variables (**X** = X_1, X_2, …, X_p) by *m* factors (**F** =f_1, f_2, …, f_m), where *m* is considerably smaller than *p. The model can be written as*

$$\mathbf{X} = \mathbf{F} \otimes \mathbf{A}$$

where A is the matrix of factor loadings. and $\otimes$ is the boolean multiplication. For n observations (cases), the data, factor scores, and factor loadings matrices may be pictured as

$$\mathbf{X}_{\mathbf{n \times p}} = \mathbf{F}_{\mathbf{n \times m}} \otimes \mathbf{A}_{\mathbf{m \times p}}$$

where **X =(** x_{ij}) has the value zero or one.

In Boolean factor analysis, the arithmetic used in the matrix multiplication is Boolean, so the scores and loadings are binary. For example, in Boolean algebra, the result of multiplying the two vectors below is one, whereas in classical factor analysis, the result is two.

$$(1 \quad 1 \quad 0 \quad 1) \otimes \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = 1 \otimes 1 + 1 \otimes 1 + 0 \otimes 0 + 1 \otimes 0 = 1$$

In classical factor analysis the score for each case (for a particular factor) is a linear combination of all the variables: the variables with large loadings all contribute to the score. In Boolean factor analysis, a case has a score of one if it has a positive response for any of the variables dominant in the factor (those not having zero loadings) and zero otherwise. Also, in classical factor analysis it is desirable to have each variable associated with one factor (a variable should not have sizeable loadings for several factors). In Boolean factor analysis, a variable may have a loading of one for several factors.

In Boolean factor analysis, the success of the technique is measured by comparing the observed binary responses with those estimated by $\otimes$ multiplying the loadings and the scores. The method count both the negative and positive discrepancies. The positive discrepancy is the number of times the observed score is one when the analysis estimates it to be zero, and the negative discrepancy is the number of times the observed score is zero when the estimated value is one. A useful measure of agreement between the original data xij and the estimated values $\hat{\mathbf{x}}_{\mathbf{ij}}$ is the total number of discrepancies

$$\mathbf{d} = \sum_{\mathbf{i}=1}^{\mathbf{n}} \sum_{\mathbf{j}=1}^{\mathbf{p}} | \mathbf{x}_{\mathbf{ij}} - \hat{\mathbf{x}}_{\mathbf{ij}} |$$

3. RESULTS

3.1 Feasibility analysis

Fifty-four experts on different aspects of service assessment participated in the feasibility evaluation. The feasibility of eDESDE-LTC has been thoroughly described by the University of Vienna group (Zeilinger et al, 2011). Participating Countries were: Spain (n:15), Slovenia (n: 10), Austria (n:8), Bulgaria (n: 8), Norway (n:6), United Kingdom (n:3), Chile (n:2), Germany (n:1) and Italy (n: 1). DESDE-LTC fulfilled the criteria of feasibility in all four factors, with arithmetic means lower than 2.5 (best to good ratings).

Applicability obtained an arithmetic mean of 2.1. According to experts data obtained using eDESDE-LTC are very useful for further processing (e.g. health care, providing LTC). As a result of the complexity of the systems in LTC expert knowledge considered an important precondition for use. It is difficult to obtain the required information for applying the instrument.

Acceptability mean rating was 2.3. It was considered user-friendly, although its handling is not comprehensible from the beginning due to many specific terms and use of new terms that are not easy to understand without special knowledge, and more practical examples are needed

Practicality obtained the worst mean (2.4). Coding and analyses of data is quite complex and high expert knowledge is required for applying the instrument. However DESDE-LTC was rated very useful in relation to the time and effort

Relevance –related to face validity- obtained the best mean rating (1.7). According to experts, almost all aims of the project (semantic interoperability, mapping, classification) are achievable using this instrument.

3.2 Consistency

Ontology analysis

The work done is a model based on the observation and research. Its objective is to get an instrument to evaluate and differentiate services and features. The work contains the meaningful domain terms. It facilitates a narrative description of its meaning (glossary). All the work done is reusable to formalize an ontology based on the specification (Scope and Purpose) and the conceptualization activity.

The formal ontology analysis allowed for the development of a formal decimal classification of the services for long term care. The final system reaches this objective in four different ways (Annex I):

- A hierarchical and formal classification of LTC services with **89** decimal numeric codes (DESDE-LTC Classification). This classification provides meaning to terms through the way they are related to others. The classification is reusable to formalize an ontology.
- A label listing which uses the specific coding of the eDESDE-LTC instrument (DESDE-LTC Code). It combines name and number of DESDE-LTC instrument branches to provide a standard description of LTC services. Every label corresponds to a decimal numeric code.
- A standard descriptor for every DESDE-LTC code (DESDE-LTC Coding List) that summarizes main characteristics of LTC services. It allows a quick search of branches definitions.

- A standard glossary of terms. It compiles an alphabetical list of definitions of key concepts that appear on DESDE-LTC Instrument.

Structural consistency (structural validity)

Boolean factorial analysis were run at three levels (levels 0, 1 and 3) (see table 7). The majority of codes where explained by a single factor. This indicates that codes are well defined and make a consistent structure within the instrument. This analysis confirmed that main branches and secondary of eDESDE-LTC are made by codes or items that measure independent characteristics of the services being assessed.

3.3 Reliability

Table 1 shows results for the agreement between observers on the main type of care of the service attending to the presence or absence of main branches (A, D, I O, R, S). The reliability coefficient for the main branch is K = 0.9674, where 1 is the highest value, with a confidence interval of CI = (0.9362 ; 0.9987)

Table 1. Inter-observer reliability: Agreement on Main branches coding by Alpha and Beta raters

		BETA						Total
		A	**D**	**I**	**O**	**R**	**S**	
ALPHA	**A**	14	0	0	0	0	0	14
	D	0	51	0	0	0	0	51
	I	0	0	2	0	0	0	2
	O	2	0	0	39	0	0	41
	R	0	0	0	0	59	0	59
	S	0	2	0	0	0	1	3
Total		16	53	2	39	59	1	**170**

For the main branch (main type of care of the service) Alpha and Beta rates had 14 coincident appearances in “A” (Accessibility), 51 in “D” (Day care), 2 in “I” (Information), 39 in “O” (Outpatient), 59 in “R” (Residential) and 1 in “S” (Self-help/volunteer).

Table 2 shows main results according to Generalizability theory and table 3 estimates reliability starting from a different condition, in this case modifying the number of evaluators.

A reliability coefficient or index of dependability of 0.96 has been found where 1 is the highest value of the coefficient. Given the hypothetical case of information gathered by one rater, the level of reliability would also be high (0.9322). The value of the coefficient rises as the number of judges increases.

Table 2. DESDE-LTC reliability (G Theory)

Source of variance	**Differ-entiation variance**	Source of variance	**Relative error variance**	% relative	**Absolute error variance**	% absolute
	.....	E	.....		0.00127	1.7
S	2.02210		.....		.....	
	.....	ES	0.07226	100.0	0.07226	98.3
Sum of variances	**2.02210**		**0.07226**	100%	**0.07353**	100%
Standard deviation	1.42201		Relative SE: 0.26881		Absolute SE: 0.27116	
Coef_G relative	0.97					
Coef_G absolute	**0.96**					

When assessing presence or absence of I, A, S, O, D, R branches as main type of care of services, reliability coefficients are really strong, all over 0.9 for Kappa and Generalizability, except for Self-Help and Volunteer care ‘S’ where coefficients (Kappa 0.49) (Generalizability 0.66)

Table 3. Reliability modifying number of raters

	G-study		**Option 1**		**Option 2**		**Option 3**	
	Lev.	Univ.	Lev.	Univ.	Lev.	Univ.	Lev.	Univ.
E	2	INF	**1**	INF	**3**	INF	**4**	INF
S	170	INF	170	INF	170	INF	170	INF
Observac.	340		170		510		680	
Coef_G rel.	0.96550		0.93330		0.97673		0.98245	
Coef_G abs.	**0.96491**		**0.93220**		**0.97633**		**0.98214**	

3.3.1 Reliability of DESDE-LTC codes

Table 4 shows the results for the reliability study in main branches, primary branches and final branches. Inter-rater reliability of final branches was calculated for 36 codings of MTCs.

Analysis for main branches (I, A, S, O, D, R) was based on the Generalizability Theory because sub-branches are not exclusive, that is, in one service different sub-branches can appear together (ex. A1 (accessibility to communication) and A2 (accessibility to physical mobility) in an association) and Kappa needs incompatible elements to analyze.

For each main branch the primary subdivision or sub-branch (A1, A2, D0, D9 etc.) was taken into account and coded with ‘0’ or ‘1’ when absent or present. The sub-branch will be present when it is present in any of its subdivisions (I2.1.1, R9.1 etc.). The reliability in this analysis is nearly perfect for all the branches even though it is smaller than the one assessed for presence or absence of main branch due to the number of new elements incorporated in the study (sub-branches).

Agreement was strong (Kappa 0.61-0.8) for ‘accessibility to care- communication and physical mobility’ (A1, A2), ‘outpatient acute non-mobile health related care’ (O3.1), ‘self-help an volunteer care with non professional staff for accessibility to care’ (S1.2), ‘low intensity social and culture structured care’ (D8.3) and ‘residential with daily support’ (R12). The agreement was nearly perfect for the rest of the DESDE-LTC codes except for 5 codes with low levels of concordance.

There was no agreement for ‘non interactive information’ (I2.2), ‘self-help an volunteer care with non professional staff for outpatient care’ (S1.3) and ‘self-help an volunteer care with professional staff for accessibility to care’ (S2.2), as only one of the judges considered it; no agreement either for ‘outpatient home and mobile (non acute) care, related to health, 3 to 6 days a week’ (O5.1.1) and ‘outpatient home and mobile (non acute) care, not related to health, 3 to 6 days a week’ (O5.2.1) as for the same services raters considered different codes.

Table 4 DESDE-LTC inter-rater reliability : “main types of care” (MTC) in main and final branches (Kappa) (G Theory) (*n*= 435)

DESDE-LTC CODES (MTC)	*n (Alfa+ Beta)*	Kappa (*k*) *	G Theory *

Information for Care (I)	15	-	Coef_G absolute: 0.95
I1	2		
I1.1	2	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
I1.2	-	-	-
I1.3	-	-	-
I1.4	-	-	-
I1.5	-	-	-
I2	11	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
I2.1	6	*K:* 0.79 (0.40-1.00)	Coef_G absolute: 0.89
I2.1.1	-	-	-
I2.1.2	2	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
I2.2	1	-	-
Accessibility to Care (A)	17	-	Coef_G absolute: 0.97
A1	2	*K:* 0.66 (0.04-1.00)	Coef_G absolute: 0.80
A2	2	*K:* 0.66 (0.04-1.00)	Coef_G absolute: 0.80
A3	-	-	-
A4	12	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
A5	1	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
Self-help and Volunteer care (S)	9	*K:* 0.49 (-0.10-1.00)	Coef_G absolute: 0.66
S1	6		
S1.1	-	-	-
S1.2	5	*K:* 0.79 (0.40-1.00)	Coef_G absolute: 0.89
S1.3	1	-	-
S1.4	-	-	-
S1.5	-	-	-
S2	3		
S2.1	2	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
S2.2	1	-	-
S2.3	-	-	-
S2.4	-	-	-
S2.5	-	-	-
Outpatient Care (O)	120	-	Coef_G absolute: 0.95
O1	-	-	-
O1.1	-	-	-
O1.2	-	-	-
O2	2		
O2.1	2	-	-
O2.2	-	-	-
O3	19	-	-
O3.1	18	*K:* 0.64 (0.38-1.00)	Coef_G absolute: 0.79
O3.2	-		
O4	-		
O4.1	-		
O4.2	-		
O5	26		
O5.1	10	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
O5.1.1	6	-	-
O5.1.2	-	-	-
O5.1.3	-	-	-
O5.2	1	-	-
O5.2.1	7	-	-
O5.2.2	-	-	-
O5.2.3	2	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
O6	6		
O6.1	4	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
O6.2	2	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
O7	-	-	-
O7.1	-	-	-
O7.2	-	-	-
O8	22		
O8.1	22	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
O8.2	-	-	-
O9	35		
O9.1	35	*K:* 0.96 (0.90-1.00)	Coef_G absolute: 0.98
O9.2	-	-	-
O10	10		
O10.1	10	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
O10.2	-	-	-
Day Care (D)	129	-	Coef_G absolute: 0.97
D0	-		
D0.1	-	-	-

D0.2	-	-	-
D1	27		
D1.1	-	-	-
D1.2	27	*K:* 0.95 (0.88-1.00)	Coef_G absolute: 0.98
D2	8		
D2.1	-	-	-
D2.2	8	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
D3	16		
D3.1	-	-	-
D3.2	16	*K:* 0.93 (0.79-1.00)	Coef_G absolute: 0.96
D4	74		
D4.1	42	*K:* 0.97 (0.92-1.00)	Coef_G absolute: 0.99
D4.2	4	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
D4.3	28	*K:* 0.92 (0.81-1.00)	Coef_G absolute: 0.98
D4.4	-	-	-
D5	-	-	-
D6	-	-	-
D6.1	-	-	-
D6.2	-	-	-
D7	-	-	-
D7.1	-	-	-
D7.2	-	-	-
D8	4		
D8.1	-	-	-
D8.2	-	-	-
D8.3	4	*K:* 0.79 (0.40-1.00)	Coef_G absolute: 0.89
D8.4	-	-	-
Residential Care (R)	126	-	Coef_G absolute: 0.99
R0	-	-	-
R1	-	-	-
R2	20	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
R3	-		
R3.0		-	-
R3.1		-	-
R3.1.1		-	-
R3.1.2		-	-
R4	21	*K:* 0.84 (0.67-1.00)	Coef_G absolute: 0.92
R5	15	*K:* 0.93 (0.79-1.00)	Coef_G absolute: 0.96
R6	14	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
R7	-	-	-
R8	4		
R8.1	-	-	-
R8.2	4	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
R9	6	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
R9.1	-	-	-
R9.2	-	-	-
R10			
R10.1	-	-	-
R10.2	-	-	-
R11	27	*K:* 0.95 (0.88-1.00)	Coef_G absolute: 0.98
R12	5	*K:* 0.79 (0.40-1.00)	Coef_G absolute: 0.89
R13	14	*K:* 1.00 (1.00-1.00)	Coef_G absolute: 1.00
R14	-	-	-

3.4 Validity

3.4.1 Descriptive analysis the branches of instrument

DESDE-LTC was used to evaluate and classify 1339 services representative of different European countries and regions of Spain. Information on the diagnostic group covered by the service was also given. (Table 5)

Table 5 Type of service and setting

	SPAIN	OTHER	**TOTAL**
MENTAL DISORDER	1269	6	1275
PHYSICAL DISABILITY	4	2	6
INTELL.DISABILITY.- DEVELOP.DIS..	18	5	23
ELDERLY	9	1	10
NON SPECIFIC	15	6	21
OTHER	4	0	4
TOTAL	1319	20	1339

Regarding the information gathered in Spain table 6 shows the distribution of the 1319 services studied by regions and diagnostic groups.

Table 6 Distribution of services in Spain

REGION Province	DIAGNOSTIC GROUP						
Andalucía	PHY.DIS	INT.DIS-DEV.DIS	ELDERLY	NON SPE.	OTHER	MENTAL D.	TOTAL
Almería						21	21
Cádiz						41	41
Córdoba						21	21
Granada						32	32
Huelva						17	17
Jaén						23	23
Málaga						44	44
Sevilla						52	52
Total						251	251
Cantabria							
Santander			1			20	21
Total			1			20	21
Castilla-La Mancha							
Albacete						23	23
Ciudad Real						24	24
Cuenca						11	11
Guadalajara						11	11
Toledo						27	27
Total						96	96
Cataluña							
Barcelona		3				367	370
Girona		1				40	41
Lleida		0				35	35
Tarragona		2				37	39
Total		6				479	485
Islas Baleares							

Islas Baleares		12	1			37	50
Total		12	1			37	50
Madrid							
Madrid	4		3	15	4	313	339
Total	4		3	15	4	313	339
Murcia							
Murcia			2			43	45
Total			2			43	45
Navarra							
Navarra			2			30	32
Total			2			30	32
TOTAL	4	18	9	15	4	1269	1319

According to the branches structure of the instrument, 4 levels of analysis were established:

Level 0: composed by the main 6 branches 'I', 'A', 'S', 'O', 'D' and 'R'.

Level 1: primary branches into which the main branches are divided. Number of codes for this level is 42.

Level 2: intermediate level where some of the level 1 branches reach their final division and where others still subdivide in a new branch.

Level 3: final subdivision of branches were a total of 89 codes explain each branch to its final characteristics. (Table 7)

Table 7 Primary branches division

Level 0	Branch **I**	Branch **A**	Branch **S**	Branch **O**	Branch **D**	Branch **R**	Total
Level 1	2	5	2	10	9	14	42
Level 3	8	5	10	24	22	20	89

To evaluate presence (1) or absence (0) of the main branch all the subdivisions have been analyzed, therefore, the main branch (A, I, S etc.) will be present (1) as long as one of the secondary branches is present (A1, I2.1.1 etc.). The starting point was level 3 where final characteristics of the codes are represented, then level 2, 1 and 0. 45 of the 89 codes considered in level 3 do not appear in any of the 1339 services evaluated.

To analyze the overall validity it is intended to examine the underlying dimensions of the instrument, that is, studying to what extent the codes of the instrument evaluate independent characteristics (of a service) or on the other hand are redundant and therefore disposable.

Level 0: it shows a matrix of 8034 values, 6439 ‘0’ and 1595 ‘1’. It was not possible to explain the 6 branches with a number of factors smaller than 6.

Level 1: it shows a matrix of 41509 values, 39872 ‘0’ and 1637 ‘1’. A 17 factor model adjusts 95% of the positive discrepancies, but this model does not explain 11 codes (I1, A5, O6, O10, D2, R5, R6, R8, R9, R10 and R12) being the ‘R’ branch the worst adjusted of all. Using a 23 factor model still 6 codes with prevalence lower than 5 are not explained, the percentage of positive discrepancies is 0,9% .Finally a 29 factor model explains the totality of the codes and shows 0% of positive discrepancy. Association between O3 and R2 remains constant for all the models; a new connection appears between I1 and A5

Level 3: it shows a matrix of 56238 values, 54595 ‘0’ and 1643 ‘1’. To explain more than 95% of positives it is needed a 24 factor model. 12 codes remain unexplained. Again the worst adjusted is branch ‘R’. A 29 factor model explains all codes with a prevalence higher than 5 except for D4.2. Table 8 summarizes the associations found in level 3 with a 29 factor model.

Table 8. Codes associations in a 29 factor model

SUMMARY:				
FACTOR 1	I11	I12	I22	A5
FACTOR 2	D41			
FACTOR 3	I212			
FACTOR 4	A1			
FACTOR 5	A2			
FACTOR 6	A4			
FACTOR 7	I11	S12	D83	
FACTOR 8	O91			
FACTOR 9	O521	S13		
FACTOR 10	S21			
FACTOR 11	S22			
FACTOR 12	O81			
FACTOR 13	O511			
FACTOR 14	0523			
FACTOR 15	R13			
FACTOR 16	D12			
FACTOR 17	R11			
FACTOR 18	O31	R2		
FACTOR 19	R4			
FACTOR 20	O101			
FACTOR 21	I211			
FACTOR 22	D43			
FACTOR 23	D11			
FACTOR 24	D32			
FACTOR 25	D22			
FACTOR 26	R6			
FACTOR 27	O61			
FACTOR 28	R102			
FACTOR 29	R5			

In factor 1, codes I1.1, I1.2, I2.2 and A5 are associated; this is mainly explained by the low prevalence of these codes in the data base, 1 time for I1.2, I2.2 and A5 and 3 for I1.1, which probably describes a very particular type of care of one service. In factor 7, codes I1.1, S1.2 and D8.3 appear together which can be explained basically in the same way than before, nonetheless, low intensity social and culture structured care (D8.3) is commonly but not necessarily associated to volunteer care (non-professional staff-accessibility to care S1.2). Factor 9 shows codes O5.2.1 and S1.3 connected which again is explained by low prevalence (1 appearance). Finally in factor 18, codes O3.1 and R2 appear together 114 times (100% of appearances) which indicates a very strong connection; facilities with hospital acute care (R2) usually offer outpatient acute care (O3.1) too and this is the case for the services collected in the data base but for example in Catalonia this type of care is described independently in some general hospitals.

4. CONCLUSION

eDESDE-LTC showed a high feasibility in its four domains: applicability, acceptability, practicality and relevance. It important to note that previous expertise on the ESMS/DESDE system had a notorious influence on the assessment of feasibility. Every feasibility-dimension was better rated from participants with ESMS/DESDE experience, particularly acceptability and practicality. There were no major differences across countries in the rating of the practicality while significant differences were identified in the assessment of the acceptability, practicality and relevance of three DESDE-LTC system.

The ontology analysis has allowed for the development of a decimal classification of LTC services based on ‘Main Types of Care’ (MTCs). It is accompanied with a formal description and identification labels at the eDESDE-LTC instrument. The ontology analysis has allowed for the development of a decimal classification (Annex I). It is accompanied with a formal description and identification labels at the eDESDE-LTC instrument. DESDE-LTC is a system focused on the standardised description and classification of services for Long-Term Care (LTC) in Europe which has a high semantic interoperability and can be used in different information systems in this region.

Structural consistency is adequate according to the factor Boolean analysis. The eDESDE-LTC codes are well defined and make a consistent structure within the instrument. This analysis confirmed that main branches and secondary of eDESDE-LTC are made by codes or items that measure independent characteristics of the services being assessed.

The external reliability obtained a high inter-observer agreement. DESDE-LTC showed high inter-rater reliability for main branches. Reliability was also high for final branches which correspond to MTCs. The branches with lower inter-observer agreement where some Information and self-support codes and special forms of outpatient mobile care. These results are better than those of the parent instruments (ESMS/DESDE)

(Salvador-Carulla et al, 2000, Salvador-Carulla et al, 2006), mainly due to the improvement of the training system which has added an online training toolkit, ant to a better formalisation of the service assessment instrument and its coding system. Descriptive validity and the structural analysis of the system were appropriate.

5. REFERENCES

Andrews G, Peters L, Teeson M (1994). Measurement of consumer outcome in mental health: A report to the National Mental Health Information Strategies Committee. Sydney: Clinical Research Unit for Anxiety Disorders.

Chisholm D, Knapp MR, Knudsen HC, Amaddeo F, Gaite L, van Wijngaarden Client Socio-Demographic and Service Receipt Inventory--European Version: development of an instrument for international research. EPSILON Study 5. European Psychiatric Services: Inputs Linked to Outcome Domains and Needs. Br J Psychiatry Suppl. 2000;(39):s28-33.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Pschicological Measurement, 20(1), 37-46

Cronbach B, Gleser L.J, Nanda, H., & Rajaratnam, N. (1972). The dependability of behavioral measurements: Theory of generalizability for scores and profiles. New York: John Wiley.

Furr, M. & Bacharach, V.R. (2008): Psychometrics: An Introduction. Thousand Oaks, CA: SAGE Publications.

González-Gordon R.G., Salvador-Carulla L, Romero C, González-Saiz F, Romero D. (2002). Feasibility, reliability and validity of the Spanish version of Psychiatric Assessment Schedule for Adults with Developmental Disability: a structured psychiatric interview for intellectual disability. J Intell Disabil Res; 46(3): 209-217

Gruber TR (1993) Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal Human-Computer Studies 43:.907-928.

Hirdes JP, Smith TF, Rabinowitz T, Yamauchi K, Pérez E, Telegdi NC, Prendergast P, Morris JN, Ikegami N, Phillips CD, Fries BE; Resident Assessment Instrument-Mental Health Group. The Resident Assessment Instrument-Mental Health (RAI-MH): inter-rater reliability and convergent validity. J Behav Health Serv Res. 2002;29(4):419-32.

International Health Facility Assessment Network (IHFAN): Profiles of Health Facility Assessment Methods. Measure Evaluation, US-Aid and WHO, 2008.

Ishak, WW; Burt, T.; Sederer, Ll. (eds)(2002) Outcome Measurement in Psychiatry: A Critical Review. Washington, D.C., American Psychiatric Publishing, Inc.

Judd, C. M., Smith, E. R. and Kidder, L. H. (1991). Research methods in social relations (6th ed.). Fort Worth, TX: Holt, Rinehart and Winston.

Keprt, A., Snàse, V., Binary Factor Analysis with Help of Formal Concepts Ostrava, Czech Republic, 2004; pp. 90-101

Kramer MS, Feinstein AR (1981). Clinical biostatistics: LIV. The biostatistics of concordance. Clinical and Pharmacological Therapy, 29, 111-123

Lloyd-Evans B, Johnson S, Slade M. Assessing the content of mental health services: a review of measures. Soc Psychiatry Psychiatr Epidemiol. 2007;42(8):673-82.

Romá-Ferri MT, Palomar M. (2008) [Analysis of health terminologies for use as ontologies in healthcare information systems]. Gac Sanit.; 22(5):421-33.

Romá-Ferri, MT. "OntoFIS: Tecnología ontológica en el dominio farmacoterapéutico". Tesis Doctoral. Departamento de Lenguajes y Sistemas Informáticos. Universidad de Alicante, 2009.

Romá-Ferri, M.T .; Palomar, M. "Interoperabilidad semántica de ontologías basada en técnicas de procesamiento del lenguaje natural (PLN)". En: Gascón, J.; Burguillos, F.; Pons, A. "La dimensión humana de la organización del conocimiento". 7º Congreso International Society for Konwledge Organization (ISKO)-España. Barcelona: Publicaciones de la Universidad de Barcelona. Departamento de Biblioteconomía y Documentación, 2005: 534-548.

Salvador-Carulla L, Romero C, Martinez A, Haro JM, Bustillo G, Ferreira A, Gaite L, Johnson S; PSICOST Group. Assessment instruments: standardization of the European Service Mapping Schedule (ESMS) in Spain. Acta Psychiatr Scand Suppl. 2000; 405:24-32.

Salvador-Carulla L, Poole M, Gonzalez-Caballero JL, Romero C, Salinas JA, Lagares-Franco CM. disabilities (DESDE) (2006).Development and usefulness of an instrument for the standard description and comparison of services for disabilities(DESDE). Acta Psychiatr Scand: 114 (Suppl. 432): 1–11

Seyrlehner D Evaluating the DESDE-LTC Instrument on feasibility; Master Thesis in Psychology (supervisor: Prof. Germain Weber – Dir); University of Vienna, 2010.

Slade M, Thornicroft G, Glover G (1999). The feasibility of routine outcome measurement. Social Psychiatry and Psychiatric Epidemiology, 34: 243-249.

Steel P, Schmidt J, Shultz J. (2008). Refining the relationship between personality and subjective well-being. Psychol Bull; 134(1):138-61.

Valls A, Gibert K, Sánchez D, Batet M. Using ontologies for structuring organizational knowledge in Home Care assistance. Int J Med Inform. 2010;79(5):370-87

Zeilinger, E., Brehmer, B. Seyrlehner, D. and Weber G. for the eDESDE-LTC Group. eDESDE-LTC: Quality Assessment and Evaluation Package [Internet]. Jerez (Spain): PSICOST and Telnet; 2011. Available from: http://www.edesdeproject.eu