DEVELOPMENT OF QUALITY INDICATORS FOR MENTAL HEALTHCARE IN THE DANUBE REGION

Isabell Lehmann[1], Dan Chisholm[2], Hristo Hinkov[3], Cyril Höschl[4], Gábor Kapócs[5,6], Tamás Kurimay[5], Dusica Lecic-Tosevski[7], Vladimir Nakov[3], János M. Réthelyi[8], Petr Winkler[9], Jürgen Zielasek[1] & Wolfgang Gaebel[1,10]

[1]*LVR-Institute for Healthcare Research, Cologne, Germany*
[2]*WHO Regional Office for Europe, Copenhagen, Denmark*
[3]*NCPHA-National Center of Public Health and Analyses, Sofia, Bulgaria*
[4]*National Institute of Mental Health, Klecany, Czech Republic*
[5]*Buda Family Centred Mental Health Centre, Department of Psychiatry and Psychiatric Rehabilitation, Teaching Department of Semmelweis University, Saint John Hospital, Budapest, Hungary*
[6]*Institut for Behavioral Sciences, Semmelweis University, Budapest, Hungary*
[7]*Serbian Academy of Sciences and Arts, Institute of Mental Health, Belgrade, Serbia*
[8]*Department of Psychiatry and Psychotherapy, Semmelweis University, Budapest, Hungary*
[9]*Department of Social Psychiatry, National Institute of Mental Health, Klecany, Czech Republic*
[10]*LVR-Klinikum, Department of Psychiatry, Medical Faculty, Heinrich-Heine-University, Düsseldorf, Germany*

SUMMARY

Background: Quality indicators are quality assurance instruments for the evaluation of mental healthcare systems. Quality indicators can be used to measure the effectiveness of mental healthcare structure and process reforms. This project aims to develop quality indicators for mental healthcare systems in Bulgaria, the Czech Republic, Hungary and Serbia to provide monitoring instruments for the transformation of mental healthcare systems in these countries.

Methods: Quality indicators for mental healthcare systems were developed in a systematic, multidisciplinary approach. A systematic literature study was conducted to identify quality indicators that are used internationally in mental healthcare. Retrieved quality indicators were systematically selected by means of defined inclusion and exclusion criteria. Quality indicators were subsequently rated in a two-stage Delphi study for relevance, validity and feasibility (data availability and data collection effort). The Delphi panel included 22 individuals in the first round, and 18 individuals in the second and final round.

Results: Overall, mental healthcare quality indicators were rated higher in relevance than in validity (Mean relevance=7.6, SD=0.8; Mean validity=7.1, SD=0.7). There was no statistically significant difference in scores between the four countries for relevance ($X^2(3)=3.581$, p=0.310) and validity ($X^2(3)=1.145$, p=0.766). For data availability, the appraisal of "YES" (data are available) ranged from 6% for "assisted housing" to 94% for "total beds for mental healthcare per 100,000 population" and "availability of mental health service facilities".

Conclusion: Quality indicators were developed in a systematic and multidisciplinary development process. There was a broad consensus among mental healthcare experts from the participating countries in terms of relevance and validity of the proposed quality indicators. In a next step, the feasibility of these twenty-two indicators will be evaluated in a pilot study in the participating countries.

Key words: *mental healthcare, psychiatry - quality indicator - quality assurance - quality management - health system performance - Danube region*

* * * * *

INTRODUCTION

Mental disorders are among the most prevalent and disabling disorders with an estimate of affecting more than a third of the European population each year (WHO Regional Office for Europe 2013, Wittchen et al. 2011). Lack of psychiatrists, regional access barriers to mental healthcare, reducing stigma and discrimination as well as the need for deinstitutionalization and adequate quality assurance measures are currently important issues in all European countries (WHO Regional Office for Europe 2013).

There is an increasing pressure on national healthcare systems to provide timely, safe, effective and high quality health care, including mental healthcare. Reform of mental healthcare systems towards achieving these goals is a central joint theme in countries of the Danube region including Bulgaria, the Czech Republic, Hungary and Serbia.

Quality indicators are widely accepted as important instruments in quality assurance, management and development (Ovretveit 2005, Gaebel et al. 2015). They can be used to increase transparency of mental healthcare, to assess the current status of mental healthcare quality and to monitor the effects of reform processes. Moreover, they can be used to compare and benchmark national mental healthcare systems. Quality indicators usually examine the structures, processes and outcomes on

Isabell Lehmann, Dan Chisholm, Hristo Hinkov, Cyril Höschl, Gábor Kapócs, Tamás Kurimay, Dusica Lecic-Tosevski, Vladimir Nakov,
János M. Réthelyi, Petr Winkler, Jürgen Zielasek & Wolfgang Gaebel: DEVELOPMENT OF QUALITY INDICATORS FOR MENTAL
HEALTHCARE IN THE DANUBE REGION   Psychiatria Danubina, 2018; Vol. 30, No. 2, pp 197-206

different healthcare levels: the national healthcare system (macro-level), mental healthcare institutions (meso-level), individual healthcare professionals and patients (micro-level) (Donabedian 1988, Gaebel et al. 2015).

Quality indicators are being developed by various stakeholders, including professional organizations, patient and care giver representatives and political organizations on regional or national levels, such as in Germany (Großimlinghaus et al. 2013, Weinmann & Becker 2009) or on an international level, (e.g., Fisher et al. 2013, Gaebel et al. 2012, Hermann & Mattke 2004, Jordans et al. 2016). However, quality indicators for countries in the Danube region are still lacking, although they may provide valuable information on the current mental healthcare reform processes in these countries.

This study therefore addresses the systematic development of mental healthcare quality indicators for Bulgaria, the Czech Republic, Hungary and Serbia. It identifies quality domains in mental healthcare as the basis for quality indicators to monitor mental healthcare reforms and ultimately improve quality of mental healthcare. Moreover, quality indicators are suggested that are considered relevant, valid and feasible by an international panel of mental healthcare experts from these countries. This paper gives an overview of the quality indicator development process and its results.

METHODS

Quality indicators were developed in a systematic, multi-step process (Figure 1). This process included a systematic literature search and a two-phase Delphi study with the participation of relevant stakeholders from Bulgaria, the Czech Republic, Hungary and Serbia.
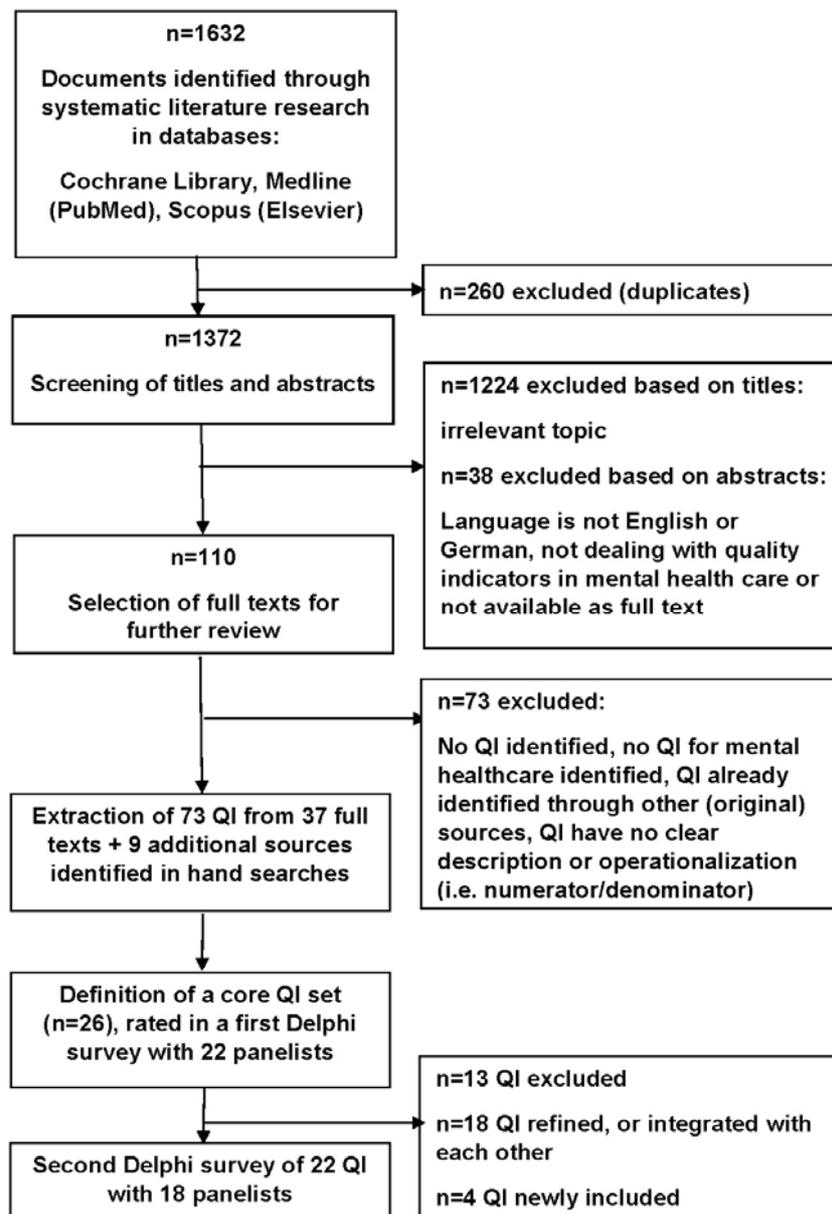

Figure 1. Overview of QI development process

Systematic literature search

In order to identify existing quality indicators for mental healthcare, a systematic literature search was conducted in the databases Pubmed, Cochrane Library and Scopus (date of search: 7 April 2017) applying the search terms ["psychiatr*" OR "mental health" OR "mental healthcare" OR "mental health care"] AND ["quality indicator*" OR "quality measure*" OR "quality assessment*" OR "quality of care measure*" OR "quality of healthcare measure*" OR "performance indicator*" OR "performance measure*"] with the asterisk indicating a truncation. The selection process of relevant articles is shown in Figure 1. The following inclusion and exclusion criteria were applied:

Inclusion criteria:
- Publication deals with quality indicators (QI) in mental healthcare;
- QI is diagnostically generic;
- QI can be applied on macro-(system-)level and meso-(institution-)level in mental healthcare.

Exclusion criteria:
- QI is specifically diagnosis-related;
- QI with a focus on specific patient groups (e.g., only children and adolescents, homeless people, women or persons in forensic psychiatry;
- QI includes very country-specific structures or institutions and cannot be generalized to other countries;
- QI without operationalization (e.g., numerator and denominator not shown).

The quality indicators identified in this systematic search were categorized into different quality domains by the project group (Table 1). The group consensually focused on the following quality domains to be preferably included in a comprehensive QI-Set:
- Mental health policies and legislation;
- Financing and costs of mental healthcare;
- Availability, accessibility and utilization of mental healthcare;
- Mental health reporting and monitoring;
- Continuity, coordination and cooperation;
- Workforce in mental healthcare;
- Promotion of mental health, and preventing mental disorders, stigma and discrimination;
- Integration of research and innovation;
- Recovery, participation and integration of persons with mental disorders;
- Patient safety.

Quality indicators were extracted from the literature (n=73) and integrated into an inventory, differentiated by the quality domains. Their number was further reduced by excluding indicators "with only minor variations to other indicators" (e.g., if several QI covered the same topic, one "key indicator" was defined among these QI), or with "unspecified operationalization" in comparison to other QI. Thus a core set of 26 indicators was created to be rated in the Delphi study.

Delphi study

The Delphi method is a multi-stage survey technique with feedback after each survey (Vorgrimler & Wübben 2003). This method was chosen because it can be used to systematically identify a group opinion. Moreover, it can be administered remotely and anonymously, making it a cost-effective research approach (Hsu & Sandford 2007). A positive ethics vote for the Delphi study was acquired from the ethics committee of the Medical Faculty of the Heinrich-Heine University Düsseldorf. All participants of the Delphi panel gave their written informed consent.

In this Delphi study, aspects of the RAND-UCLA Appropriateness Method (Fitch et al. 2001) were applied. This means that two rounds of quality indicator ratings took place. In between rounds, members of the project group discussed the statistically summarized ratings and pseudonymous comments of the first Delphi round to provide an input for the second Delphi round.

Compilation of the Delphi panel

In this Delphi study, a multidisciplinary group of experts, who were knowledgeable in the field of quality in mental healthcare, was created. Experts included mental healthcare professionals such as psychiatrists, psychologists, social workers, mental healthcare managers/administrators, policy makers and patient representatives. The multidisciplinary panel was assembled with the aim to increase acceptance and to facilitate the implementation process for quality indicators. Therefore, it was considered very important to include knowledgeable individuals, who could facilitate the implementation of the quality indicators. Each country partner (H. Hinkov, Bulgaria; C. Höschl, Czech Republic; T. Kurimay, Hungary; D. Lecic-Tosevski, Serbia) nominated four to five experts from each country. In total, a list of twenty-two experts was compiled.

Development of the Delphi survey instrument

The design of the Delphi instrument drew on our previous experience in the systematic development of quality indicators in Germany (Großimlinghaus et al. 2013, Wobrock et al. 2011, Zielasek et al. 2012) as well as other international projects in which QI were developed employing systematic evidence- and consensus-based processes (Parameswaran et al. 2015, Hermann et al. 2004, Jordans et al. 2016).

The Delphi survey forms included four criteria per QI. These criteria were extracted from the German Instrument for the Assessment of Quality Indicators (QUALIFY) (Reiter et al. 2007) and were defined as follows:
- ***Relevance:*** The QI captures a topic that is or should be relevant to mental healthcare planning and monitoring with the aim to assure and improve quality of mental healthcare.

Table 1. Project group

Dan Chisholm	Project partner, WHO Regional Office for Europe, Denmark
Wolfgang Gaebel	Project leader, Düsseldorf/Cologne
Isabell Lehmann	Scientific study coordinator, Düsseldorf/Cologne
Hristo Hinkov, Vladimir Nakov	Project partners, Bulgaria
Cyril Höschl, Petr Winkler	Project partners, Czech Republic
Tamás Kurimay, Gábor Kapócs	Project partners, Hungary
Dusica Lecic-Tosevski	Project partner, Serbia
Dijana Naumoska	Project coordinator, Düsseldorf/Cologne
Jürgen Zielasek	Project partner, Düsseldorf/Cologne

▪ *Validity:* The QI is defined clearly and unambiguously.

▪ *Feasibility:*

• *Data availability:* It is likely that data needed to measure this QI are routinely collected electronically (i.e. in country-specific statistics, data documentations in mental healthcare, questionnaires) and available from a database/databases.

• *Data collection effort:* The additional effort for electronic collection of data for this QI is justifiable.

The survey instrument requested the panelists to rate relevance and validity on a 9-point-Likert scale. The feasibility criteria data availability and data collection effort were rated with "Yes", "No" or "Uncertain". In addition, the Delphi panelists could provide free text commentaries on each QI.

Rating process and analysis

There were two Delphi rounds for the rating process. After the first round, the project group reviewed the summarized results for each criterion for each QI (mean, standard deviation (SD) for relevance and validity; percentage for data availability and data collection effort) together with the pseudonymous comments for each indicator. Between rounds, the quality indicators were further refined based on comments made in the first Delphi survey. This led to the exclusion of some indicators and the inclusion of new ones (Figure 1). All first-round results were summarized and provided to the Delphi panelists, so that they could review their own opinions and ratings in the context of the combined summarized ratings and pseudonymous comments. In the final rating round, a total of 22 indicators from eight quality domains were rated by 18 panelists (Figure 1).

Descriptive univariate statistics were performed for each criterion per indicator using IBM® SPSS Statistics® Version 22 to create a ranking of indicators for relevance and validity. Explorative analyses by means of nonparametric tests were performed to further analyze the results. To identify differences of the relevance and validity scores between countries, Kruskal-Wallis tests were performed.

RESULTS

The following results were obtained after the second and final Delphi survey. Eighteen experts participated (Bulgaria n=5; the Czech Republic n=4; Hungary n=4; Serbia n=4; WHO Regional Office for Europe n=1). Different professions were involved, including one service user participant (Table 2).

Table 2. Characteristics of participants of the Delphi survey (n=18)

	Number of participants
Country	
Bulgaria	5
Czech Republic	4
Hungary	4
Serbia	4
Other	1
Gender	
Male	11 (61%)
Female	7 (39%)
Age	
25-34	2 (11%)
35-44	2 (11%)
45-59	11 (61%)
60 and older	3 (17%)
Participant's profession/perspective	
Government official	5 (28%)
Mental healthcare manager – inpatient care	6 (33%)
Mental healthcare manager – outpatient care	4 (22%)
Mental healthcare manager – community care	2 (11%)
Mental healthcare manager – other	3 (17%)
Mental health professional - psychiatrist	10 (56%)
Mental health professional - psychologist	1 (6%)
Mental health professional - other	4 (22%)
Mental healthcare researcher	12 (67%)
Service user representative	1 (6%)

Table 3 shows the summarized results of all ratings ranked by relevance. Regarding the ten most highly ranked indicators for relevance, the indicators were rated between 8.2 (mean; SD = 0.7) and 7.7 (mean; SD = 1.2). Table 4 summarizes the results ranked for validity. Here, the ten most highly ranked indicators were rated between 7.9 (mean; SD = 0.8) and 6.1 (mean; SD = 1.8).

The range of relevance ratings for all 22 indicators was between 6.4 and 8.2 and the range for validity was 6.1 to 7.9. Overall, ratings for relevance were higher than for validity. The overall mean for relevance was 7.6 (SD=0.8) and of validity 7.1 (SD=0.7) (Wilcoxon signed-rank test, z=-2.936, p=0.003).

Seven indicators were included in both top-ten relevance and validity rankings:
- "Utilization and coverage of mental health services (bipolar disorder and schizophrenia)",
- "Availability of mental health service facilities",
- "Mental health legislation",
- "Availability and content of a mental health action plan document",
- "Utilization and coverage of mental health services (anxiety disorders and depression)",
- "Utilization and coverage of mental health services (alcohol use disorder)", and
- "Involuntary inpatient admissions".

Table 3. Ranking of quality indicators (QI) according to "relevance" (n=22)

Rank	QI	Relevance: Mean (SD), Range	Validity: Mean (SD), Range	Data availability: "YES" in %
1	Utilization and coverage of mental health services (bipolar disorder and schizophrenia)	8.2 (0.7), 7-9	7.8 (0.8), 7-9	56%
2	Availability of mental health service facilities	8.2 (0.6), 7-9	7.9 (0.8), 6-9	94%
3	Mental health legislation	8.1 (0.9), 6-9	7.7 (0.8), 6-9	83%
4	Availability and content of a mental health action plan document	8.0 (0.9), 6-9	7.3 (1.3), 4-9	61%
5	Utilization and coverage of mental health services (anxiety disorders and depression)	7.8 (0.9), 6-9	7.4 (0.8), 6-9	28%
6	Total beds for mental health care per 100,000 population	7.8 (1.0), 5-9	7.1 (1.2), 5-9	94%
7	Human resources in mental health facilities per capita	7.8 (1.5), 3-9	7.1 (1.2), 3-9	50%
8	Involuntary inpatient admissions	7.7 (1.1), 5-9	7.6 (1.2), 5-9	50%
9	Utilization and coverage of mental health services (alcohol use disorder)	7.7 (1.1), 5-9	7.4 (1.1), 5-9	33%
10	Health budget	7.7 (1.2), 5-9	7.2 (1.2), 4-9	33%
11	Utilization and coverage of mental health services (children and adolescents with conduct disorder)	7.6 (1.0), 5-9	7.3 (1.1), 4-9	28%
12	Utilization and coverage of mental health services (children and adolescents with intellectual disabilities)	7.6 (1.2), 5-9	7.3 (1.3), 4-9	33%
13	Utilization and coverage of mental health services (substance-use disorder, other than alcohol)	7.6 (1.3), 5-9	7.3 (1.3), 4-9	28%
14	Follow-up of visits after mental health-related hospitalization	7.6 (1.5), 4-9	7.3 (1.1), 5-9	17%
15	Utilization and coverage of mental health services (dementia)	7.5 (1.2), 5-9	7.2 (1.3), 5-9	28%
16	Equity	7.4 (1.6), 3-9	6.6 (1.9), 3-9	72%
17	Integration of care	7.4 (1.3), 5-9	6.5 (1.7), 2-9	11%
18	User associations and mental health policies, plans or legislation	7.3 (1.6), 4-9	6.4 (1.2), 5-9	22%
19	Assisted housing	7.2 (1.4), 5-9	6.4 (1.4), 4-9	6%
20	Total national expenditure on mental health services per capita per year	7.1 (1.7), 2-9	7.3 (1.2), 5-9	39%
21	Multi-disciplinary community mental health teams	6.9 (1.6), 4-9	6.5 (1.5), 3-9	22%
22	Anti-stigma movement	6.4 (1.9), 1-8	6.1 (1.8), 1-8	33%

Table 4. Ranking of quality indicators (QI) according to "validity" (n=22)

Rank	QI	Validity: Mean (SD), Range	Relevance: Mean (SD), Range	Data availability: "YES" in %
1	Availability of mental health service facilities	7.9 (0.8), 6-9	8.2 (0.6), 7-9	94%
2	Utilization and coverage of mental health services (bipolar disorder and schizophrenia)	7.8 (0.8), 7-9	8.2 (0.7), 7-9	56%
3	Mental health legislation	7.7 (0.8), 6-9	8.1 (0.9), 6-9	83%
4	Involuntary inpatient admissions	7.6 (1.2), 5-9	7.7 (1.1), 5-9	50%
5	Utilization and coverage of mental health services (anxiety disorders and depression)	7.4 (0.8), 6-9	7.8 (0.9), 6-9	28%
6	Utilization and coverage of mental health services (alcohol use disorder)	7.4 (1.1), 5-9	7.7 (1.1), 5-9	33%
7	Follow-up of visits after mental health-related hospitalization	7.3 (1.1), 5-9	7.6 (1.5), 4-9	17%
8	Utilization and coverage of mental health services (children and adolescents with conduct disorder)	7.3 (1.1), 4-9	7.6 (1.0), 5-9	28%
9	Total national expenditure on mental health services per capita per year	7.3 (1.2), 5-9	7.1 (1.7), 2-9	39%
10	Availability and content of a mental health action plan document	7.3 (1.3), 4-9	8.0 (0.9), 6-9	61%
11	Utilization and coverage of mental health services (children and adolescents with intellectual disabilities)	7.3 (1.3), 4-9	7.6 (1.2), 5-9	33%
12	Utilization and coverage of mental health services (substance-use disorder, other than alcohol)	7.3 (1.3), 4-9	7.6 (1.2), 5-9	28%
13	Health budget	7.2 (1.2), 4-9	7.7 (1.2), 5-9	33%
14	Utilization and coverage of mental health services (dementia)	7.2 (1.3), 5-9	7.5 (1.2), 5-9	28%
15	Total beds for mental health care per 100,000 population	7.1 (1.2), 5-9	7.8 (1.0), 5-9	94%
16	Human resources in mental health facilities per capita	7.1 (1.2), 3-9	7.8 (1.5), 3-9	50%
17	Equity	6.6 (1.9), 3-9	7.4 (1.6), 3-9	72%
18	Integration of care	6.5 (1.7), 2-9	7.4 (1.3), 5-9	11%
19	User associations and mental health policies, plans or legislation	6.4 (1.2), 5-9	7.3 (1.6), 4-9	22%
20	Assisted housing	6.4 (1.4), 4-9	7.2 (1.4), 5-9	6%
21	Multi-disciplinary community mental health teams	6.5 (1.5), 3-9	6.9 (1.6), 4-9	22%
22	Anti-stigma movement	6.1 (1.8), 1-8	6.4 (1.9), 1-8	33%

These seven indicators cover three quality domains: „mental health policies and legislation" (two QI), „availability, accessibility and utilization of care service structures" (four QI), and „patient safety" (one QI).

A Spearman rank correlation test was performed for all 22 indicators showing that mean validity and relevance ratings were significantly correlated (P=0.747, p<0.000). When comparing the four countries with each other, there was no statistically significant difference in relevance and validity scores (Kruskal Wallis Test analyses for relevance $X^2(3)=3.581$, p=0.310, and validity $X^2(3)=1.145$, p=0.766). The mean scores for each country are shown in Table 5.

With regard to data availability, there was a wide range for the rating "YES" (data needed to measure a QI are routinely collected electronically and available form a database) between 6% for "assisted housing" and 94%

for "total beds for mental health care per 100,000 population" as well as 94% for "availability of mental health service facilities".

The ratings for data availability of the top-ten QI ranked by relevance range from 28% for "YES" ("Utilization and coverage of mental health services (anxiety disorders and depression)") to 83% for "YES" ("Mental health legislation"). There was a statistically significant difference in his expert-estimated data availability between countries in one indicator "Utilization and coverage of mental health services (alcohol disorder)" ($X^2(8)=18.806$, p=0.016). This result was mainly due to the rating for the Czech Republic, which indicated that data are likely to be available there (Table 6). However, for the other six indicators, there were no statistically significant differences in expert-estimated data availability between countries (Table 6).

Table 5. Overall mean relevance and validity scores across all quality indicators (n=22) per country

Country	Relevance (Mean, SD)	Validity (Mean, SD)
Bulgaria	7.1 (0.8)	6.9 (0.8)
Czech Republic	7.8 (0.2)	7.5 (0.2)
Hungary	8.0 (1.2)	7.2 (0.9)
Serbia	7.6 (1.0)	7.1 (1.0)

Table 6. Differences between countries in data availability of seven quality indicators (QI), overlapping in the top-ten ranking of relevance and validity

QI	Data availability*							
	Bulgaria		Czech Republic		Hungary		Serbia	
	Yes	No	Yes	No	Yes	No	Yes	No
Mental health legislation	80%	0%	75%	0%	75%	25%	100%	0%
Availability and content of a mental health action plan document	60%	20%	100%	0%	25%	0%	50%	25%
Availability of mental health service facilities	80%	0%	100%	0%	100%	0%	100%	0%
Utilization and coverage of mental health services (bipolar disorder and schizophrenia)	40%	40%	75%	0%	75%	0%	50%	25%
Utilization and coverage of mental health services (alcohol use disorder)	0%	20%	100%	0%	50%	25%	0%	75%
Utilization and coverage of mental health services (anxiety disorders and depression)	40%	40%	50%	0%	25%	25%	0%	75%
Involuntary inpatient admissions	20%	40%	50%	25%	50%	0%	50%	50%

*Does not include ratings for "uncertain", thus Yes/No ratings per country may not add up to 100%.

DISCUSSION

The goal of this study was to develop quality indicators for mental healthcare systems in the countries of the Danube region with the aim to monitor mental healthcare reforms and to improve the quality of mental healthcare in these countries. This is the first study that developed a common set of quality indicators in these countries by means of a systematic development process. The indicators were considered both relevant and valid by expert opinion, especially regarding the seven indicators that overlap in the top-ten rankings of relevance and validity (Tables 3 & 4).

When considering the highest rated indicators for the criterion "relevance", the following quality domains are included: Mental health policies and legislation; availability, accessibility and utilization of care services; financing and costs of mental healthcare; workforce in mental healthcare and patient safety. These domains reflect mostly levels of the overall healthcare system (macro-level) and the institutional-(meso)-level.

With regard to the seven overlapping indicators in the top ten rankings of validity and relevance, the quality domains "mental health policies and legislation", "availability, accessibility and utilization of care service structures", and "patient safety" were represented. Patient safety is a domain of growing importance to be considered in quality assurance for mental healthcare. Both, safety and patient-centeredness were defined as core aspects of quality by the Organization for Eco-

nomic Co-operation and Development (OECD) (Arah et al. 2006) and safety as well as effectiveness of care should be the focus of mental healthcare reforms and included in mental health policies (WHO 2013). The development and content of mental health policies, the availability of different care structures including specialized services, and the availability of community services and teams are all topics included in this indicator set. Since the action plan of the European Joint Action on Mental health and Wellbeing on the transformation of mental healthcare systems towards deinstitutionalized community-based mental health services (Caldas Almeida et al. 2015) recommends these activities, our proposed indicators may serve as a monitoring tool of the implementation of the Joint Action's aims. Moreover, these indicators may be used as an extension of the indicators of the WHO Mental Health Atlas (2015), which provide information on the availability of mental health services and resources in different countries. Our indicators provide additional information as thay also focus on the utilization and coverage of mental health services.

There were no significant differences in ratings for relevance and validity between countries. This indicates an implicit consensus across the four countries in terms of the relevance and validity of quality indicators. All indicators were rated significantly higher in relevance than validity. This indicates that while they may cover a relevant quality aspect, their operationalizations may need improvement. This is especially relevant for the

implementation of quality indicators across countries, which requires the documentation of similar data if indicators are to be compared internationally. Thus, there is a need to field study the implementation of the indicators and to design large-scale studies, which may be able to show whether the country-wide implementation of indicators leads to measurable improvements of mental healthcare.

Our analyses for the seven indicators that overlap in both top-ten relevance and validity rankings showed that for six indicators there were no significant differences between countries as to whether data were likely to be already collected electronically and thus were likely to be available. This may provide an important prerequisite for the implementation of this QI set. However, data availability, data quality and data accessibility need to be further investigated in more detail. For one indicator ("Utilization and coverage of mental health services (alcohol disorder)") data were reported to be likely available in the Czech Republic, but not in Bulgaria, Hungary or Serbia. This may be explained by the availability of both, national registers of health care utilization and recent community-based epidemiological survey, in the Czech Republic.

The more data are systematically and electronically collected, the more likely it is that indicators may be measureable, since a high workload due to separate data collections for quality indicators may strongly decrease feasibility (Gaebel et al. 2015). However, it is unlikely that relevant indicators can only be measured through routinely collected data that originally are not intended for quality assurance purposes, but for the remuneration of mental health services (Gaebel et al. 2015). We conducted a pilot test of quality indicators for schizophrenia and depression in ten German psychiatric hospitals, which showed that of 24 indicators only 14 could be measured by means of routinely available data, including data for remuneration of inpatient mental healthcare services and additional medical documentations (Großimlinghaus et al. 2015).

One further indicator ("Formally defined minimum data set items") was only rated in the first Delphi round. Hereafter, the project group decided to remove it from the indicator set and use it as a "meta-level" indicator to be used to determine if all data for the indicators of the proposed set are being collected. For this meta-level indicator, the minimum data set still needs to be defined, which depends on the final operationalization of all quality indicators.

Furthermore, even though diagnosis-specific indicators were excluded in the systematic literature study, it became clear during this two-stage Delphi survey that differentiating some indicators according to diagnoses may provide a more detailed insight into the utilization and coverage of mental healthcare services. Therefore, indicators focusing on the latter were differentiated by diagnostic groups during the Delphi process.

Overall, the strength of this study relates to the systematic and multidisciplinary approach, which also included different stakeholders from each country. This increases the chance of later acceptance of the quality indicators. In a Delphi study, it is important to select knowledgeable and expert participants, because this determines the quality of the results generated. It is recommended to include participants who can implement the results of the Delphi study (Hsu & Sandford 2007). In our study, many participants may support implementation of the developed indicators and due to their professional and institutional backgrounds had extensive knowledge of data availability in their countries.

A limitation of this study was that indicators were not rated on a broader basis with more stakeholders from the four countries. However, the participation of 18 individuals in our final Delphi round is higher than in other international quality indicators development studies, which included expert panels of 12 persons (Parameswaran et al. 2015, Hermann et al. 2004). Moreover, the number of participants was well balanced between the participating countries with 4-5 panelists per country.

The rating criteria of the Delphi process were defined based on what we considered important in the preselection of quality indicators including their relevance, validity and feasibility. A limitation may be that we did not apply more criteria. However, in other transnational quality indicator development processes (Hermann & Mattke 2004, Parameswaran et al. 2015, WHO 2013) the number of applied criteria was similar. In general, quality indicators can fulfill many different criteria to be fully considered relevant, scientifically sound and feasible (Reiter et al. 2008). For example, the criterion "consideration of potential risks/side effects" establishes whether there are risks, such as false incentives, through the use of an indicator. When the number of psychiatric beds, for instance, is being measured and the quality goal is that the number of beds should be decreasing, it needs to be assured that the provision of community or outpatient treatment is available, accessible and of sufficient quality for persons with mental disorders. Otherwise, there would be a risk of under provision of mental healthcare services. Therefore, the results of quality indicators need to be interpreted in the context of normative goals and can provide incentives for further quality assurance activities.

Regarding validity, the literature proposes many different definitions (Reiter et al. 2008). In this study we chose a definition that focuses on "face validity", i.e. a clear and unambiguous definition of the indicator that is likely to be of "high quality" at first sight. Other definitions of validity include, e.g., "does the indicator really measure what it intends to measure", or "is there a strong evidence base to support that the indicator can lead to improved quality". Further systematic literature studies on the evidence base of each indicator were not feasible within the scope of this study.

In a follow-up study, we will pilot test all twenty-two indicators in Bulgaria, the Czech Republic, Hungary and Serbia. This process includes identifying the necessary data and data sources as well as refining the operationalization of each indicator, acquiring data, performing data analyses and examining the plausibility of the results. The challenges posed by such a transnational study include different data structures, data availability, and data quality.

CONCLUSION

This study shows the systematic and multidisciplinary development of quality indicators in mental healthcare for Bulgaria, the Czech Republic, Hungary and Serbia.

According to the Delphi process there was a consensus among these four countries regarding the relevance and validity of the proposed quality indicators. Thus, it seems that the participating countries share equal goals and interests in reforming their mental healthcare systems.

The developed quality indicators focus on the mental healthcare system and the institutional level and may be used to monitor the effects of the ongoing mental healthcare reform processes in the participating countries. Pilot testing of these quality indicators is planned to further evaluate their feasibility.

Contribution of individual authors:

Isabell Lehmann: Research idea, coordination of study, study design, literature search, data collection, statistical analysis, manuscript writing;

Dan Chisholm: Provision of transnational information on countries, expert in quality indicator development, approval of the final version;

Hristo Hinkov, Cyril Höschl, Gábor Kapócs, Tamás Kurimay, Dusica Lecic-Tosevski, Vladimir Nakov & Petr Winkler: Provision of country specific information, expert in quality indicator development, approval of the final version;

János M. Réthelyi: Expert in quality indicator development, approval of the final version;

Jürgen Zielasek; Research idea, expert in quality indicator development, approval of the final version;

Wolfgang Gaebel: Research idea, methodological advice, expert in quality indicator development, approval of the final version.

References

1. *Arah OA, Westert GP, Hurst J, Klazinga NS: A conceptual framework for the OECD Health Care Quality Indicators Project. Int J Qual Health Care 2006; 18(Suppl. 1):5–13*

2. *Caldas Almeida JM, Mateus P, Tomé, Katschnig H, Hinkov H, Sooniste I. et al. Joint Action on Mental Health and Wellbeing. Towards community-based and socially inclusive mental health care. Situation analysis and recommendations for action. Retrieved on 21 January 2018 from http://www.mentalhealthandwellbeing.eu/assets/docs/publications/WP5%20Final-20151203075843.pdf*

3. *Donabedian A: The quality of care. How can it be assessed? JAMA 1988; 260:1743–1748*

4. *Fisher CE, Spaeth-Rublee B, Pincus HA For the IIMHL Clinical Leaders Group: Developing mental health-care quality indicators: toward a common framework International Journal for Quality in Health Care 2013; 25:75–80*

5. *Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P et al.: The RAND/UCLA Appropriateness Method User's Manual. RAND UCLA Cooperation 2001. Retrieved on 18 January 2018 from https://www.rand.org/content/dam/rand/pubs/monograph_reports/2011/MR1269.pdf*

6. *Gaebel W, Becker T, Janssen B, Munk-Jorgensen P, Musalek M, Rössler W et al.: EPA guidance on the quality of mental health services. Eur Psychiatry 2012; 27:87-113*

7. *Gaebel W, Großimlinghaus I, Heun R, Janssen B, Johnson B, Kurimay T, et al.: European Psychiatric Association (EPA) guidance on quality assurance in mental healthcare. Eur Psychiatry 2015; 30:360–387*

8. *Großimlinghaus I, Falkai P, Gaebel W, Janssen B, Reich-Erkelenz D, Wobrock T et al.: Developmental process of DGPPN quality indicators. Nervenarzt 2013; 84:350–365. [Article in German]*

9. *Großimlinghaus I, Falkai P, Gaebel W, Hasan A, Jänner M, Janssen B et al.: Assessment of quality indicators with routine data. Presentation of a feasibility test in ten specialist clinics for psychiatry and psychotherapy. Nervenarzt 2015; 86:1393–1399. [Article in German]*

10. *Hermann R & Mattke S: Selecting Indicators for the Quality of Mental Health Care at the Health Systems Level in OECD Countries. OECD Health Technical Papers, No. 17, OECD Publishing. 2004. Retrieved on 3. January 2018 from: http://dx.doi.org/10.1787/388745076135*

11. *Hermann RC, Palmer H, Leff S, Shwartz M, Provost S, Chan J, et al.: Achieving Consensus Across Diverse Stakeholders on Quality Measures for Mental Healthcare. Medical Care 2004; 42:1246-1253.*

12. *Hsu C-C, Sandford BA: The Delphi Technique: Making Sense Of Consensus. Practical Assessment, Research & Evaluation, 2007; 12(10). Retrieved on 30. December 2017 from http://pareonline.net/getvn.asp?v=12&n=10*

13. *Øvretveit J: What are the advantages and limitations of different quality and safety tools for health care? 2005. Copenhagen, WHO Regional Office for Europe (Health Evidence Network report; Retrieved on 3. January 2018 from: http://www.euro.who.int/document/e87577.pdf*

14. *Parameswaran SG, Spaeth-Rublee B, Pincus HA: Measuring the Quality of Mental Health Care: Consensus Perspectives from Selected Industrialized Countries. Adm Policy Ment Health 2015; 42:288–295*

15. *Reiter A, Fischer B, Kötting J, Geraedts M, Jäckel WH, Barlag H et al.: QUALIFY: Instrument for the Assessment*

of Quality Indicators. Version 1.0 (English). Bundesgeschäftsstelle Qualitätssicherung gGmbH, 2007. Retrieved on 07.08.2015 from http://bqs-institut.de/innovationen/qualify-instrument.html

16. *Wobrock T, Reich-Erkelenz D, Janssen B, Sommerlad K, Gaebel W, Falkai P et al.: Qualitätsindikatoren in der Psychiatrie. Psychiatrie 2010; 7:179–189. [Article in German]*

17. *Zielasek J, Großimlinghaus I, Janssen B, Wobrock T, Falkai P, Reich-Erkelenz D et al.: Die Rolle von Qualitätsindikatoren in der psychiatrischen Qualitätssicherung und aktueller Stand der Entwicklung von Qualitätsindikatoren. Die Psychiatrie 2012; 9:46-52. [Article in German]*

18. *Vorgrimler D & Wüben D: Die Delphi-Methode und ihre Eignung als Prognoseinstrument. Retrieved on 30. December 2017 from: https://www.destatis.de/DE/ Publikationen/ WirtschaftStatistik/Gastbeitraege/DelphiMethode_82003.pdf? __blob=publicationFile [Article in German]*

19. *Weinmann S & Becker T: Qualitätsindikatoren für die Integrierte Versorgung von Menschen mit Schizophrenie. 2009. Psychiatrie-Verlag, Bonn*

20. *Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B et al.: The size and burden of mental disorders and other disorders of the brain in Europe 2010, European Neuropsychopharmacology 2011; 21:655–679*

21. *WHO Regional Office for Europe: European Mental Health Action Plan. 2013. Retrieved on 3 January 2018 from: http://www.euro.who.int/__data/assets/pdf_file/0004/1941 07/63wd11e_MentalHealth-3.pdf*

22. *WHO (World Health Organization): Mental Health Action Plan 2013-2020. World Health Organization, 2013. Retrieved on 12 February 2018 from http://www.who.int/ mental_health/publications/action_plan/en/*

23. *WHO (World Health Organization): Mental Health Atlas 2014. World Health Organization, 2015. Retrieved on 6 February 2018 from: http://apps.who.int/iris/bitstream/ 10665/ 178879/1/9789241565011_eng.pdf?ua=1&ua=1*

Correspondence:

Isabell Lehmann, M.Sc
LVR-Institute for Healthcare Research, LVR-Klinik Köln
Wilhelm-Griesinger Straße 23, 51109 Köln, Germany
E-mail: Isabell.Lehmann@lvr.de