Chapter 3 – Availability, content, and quality of local guidelines for the assessment of suicide attempters in university and general hospitals in the Netherlands

BAS VERWEY, JEROEN A. VAN WAARDE, IRIS A.L.M. VAN ROOIJ, GERARD GERRITSEN AND FRANS G. ZITMAN

Published in: General Hospital Psychiatry, 2006; 28: 336-342
ABSTRACT

Objective
This study was performed to investigate the availability, content, and quality of local guidelines for the assessment of suicide attempters in the Netherlands.

Method
All university and general hospitals in the Netherlands were asked to provide their local guidelines. Published national guidelines and the AGREE instrument were used to evaluate the content and quality of the local guidelines.

Results
Eighty-eight hospitals (90.7%) responded; 34 (38.6%) reported that they used local guidelines. Twenty-seven guidelines were submitted for evaluation. Most of the guidelines were more than five years old and had not been updated recently. The contents of the guidelines differed. Criteria addressing patient safety, staff attitude toward patients, re-assessment of non-alert patients, relevant stressors, involvement of significant others, and aftercare were found in less than 50% of the guidelines. Although psychiatric consultation was incorporated in almost 80%, the psychiatrist’s tasks were specified infrequently. The guidelines seldom required monitoring of staff compliance. Only in the AGREE domain ‘clarity and presentation’ the mean score was above 60% of the maximum. According to the instructions for the AGREE instrument, 10 (37.0%) of the 27 guidelines were recommended (with provisos or alterations) and one was strongly recommended for use in practice.

Conclusions
In the Netherlands, a minority of hospitals reported use of local guidelines for the assessment of suicide attempters. When available, the guidelines were mostly not based on international standards, their contents varied greatly, and their quality was unsatisfactory.
INTRODUCTION

The World Health Organization has estimated that in the year 2000 approximately one million people worldwide died from suicide, and ten to twenty times more made a suicide attempt (World Health Organization, 2000). Of those who try, at least 2% repeat the attempt and succeed within ten years, mostly within the first two years (de Moore & Robertson, 1996; Hawton et al., 1998; Hawton, Zahl, & Weatherall, 2003; Owens, Horrocks, & House, 2002; Zahl & Hawton, 2004). A significant number of suicide attempters present for treatment to emergency departments of university and general hospitals. Assessment and treatment of these patients is often complicated. Many patients are not only in a disordered psychic state, but they are often in an acute life-threatening somatic condition as well. Therefore, treatment requires a subtle interplay between psychiatrists, other medical specialists, and nurses. Unfortunately, this interplay is frequently made more complex by a negative or ambivalent attitude by emergency department staff that leads to stigmatization and lack of empathy for the patient (Roose, 2001). Nevertheless, a thorough assessment and treatment are important to prevent somatic complications, further deliberate self-harming, and completed suicide (Hickey, Hawton, Fagg, & Weitzel, 2001; Suokas & Lonnqvist, 1991). Thus, it is critical for a hospital to provide services of high quality to these patients (Hawton & Heeringen, 2000). In general, clinical practice guidelines are a potentially valuable tool to improve quality of care (Grol, 1997), and therefore it can be argued that hospitals should adopt local guidelines for the assessment and treatment of people who attempt suicide.

In the Netherlands in 1991, the Medical Scientific Council of the National Organization for Quality Assurance in Hospitals issued national clinical guidelines for the assessment of suicide attempters in general hospitals (Centraal Begeleidingsinstituut voor de Intercollegiale Toetsing, 1991). National guidelines for the management and assessment of suicide attempters have been developed in other countries as well (American Psychiatric Association, 2003; Boyce P., Carter G., Penrose-Wall J., Wilhelm K., & Goldney R., 2003; Goldberg, 1987; Hirschfeld & Russell, 1997; Isacsson & Rich, 2001; Lonnqvist & Suokas, 1992; Packman, Marlitt, Bongar, & Pennuto, 2004; Royal College of Psychiatrists, 2004; Simon, 2002). In the few studies evaluating the application of such national guidelines in hospitals, great differences have been observed in the care provided (Barr, Leitner, & Thomas, 2005; Hawton & James, 1995; Owens & House, 1994; Slinn, King, & Evans, 2001). For example, although a national guideline for assessment of suicide attempters was established in England, only 60% of hospitals implemented a policy document addressing
this issue (Slinn et al., 2001). Recently, the APA published an elaborate set of guidelines representing ‘a synthesis of current scientific knowledge and rational clinical practice on the assessment’ To evaluate the Dutch situation, we performed this descriptive study to gain insight in the availability, content, and quality of local guidelines for assessing suicide attempters in university and general hospitals.

METHODS

In 2005, we sent a short questionnaire with a self-addressed envelope to the responsible psychiatrist in all university (n = 8) and general (n = 89) hospitals in the Netherlands. Addresses of general hospitals were obtained from the Dutch Society of Hospitals (NVZ) with which all hospitals are affiliated. Four questions were asked: [1] ‘Does your hospital use a guideline for the assessment of suicide attempters?’ [2] ‘If so, since what year?’ [3] ‘From what year dates the most recent update?’ [4] ‘Has the observance of the guideline been tested?’ (yes, once/yes, regularly/no) The psychiatrist was asked to return the questionnaire together with any available guidelines. After four weeks, all non-responders were reminded of the request by telephone.

MEASUREMENTS

A. Contents of local guidelines  
The criteria to evaluate the contents of local guidelines were adopted from the guidelines of the American Psychiatric Association (American Psychiatric Association, 2003), the Royal College of Psychiatrists (Royal College of Psychiatrists, 2004), and the Dutch National Organization for Quality Assurance in Hospitals (Centraal Begeleidingsinstituut voor de Intercollegiale Toetsing, 1991). From these sources, criteria were included only if they were discussed in all three guidelines as major topics for the assessment. Topics only occurring in one of the guidelines, for example ‘training and supervision of staff’ and ‘providing education to the patient and family’, were not selected. Specifically, we assessed whether the guidelines provided instructions to:

1. address patient safety during the assessment process
2. establish and maintain a therapeutic alliance between the clinician and the patient
promptly assess the physical condition of the patient, including the patient’s level of consciousness

perform a psychiatric consultation for all patients, specifically to:

4.1 assess suicidality
4.2 perform a psychiatric examination to detect mental illness, alcohol abuse, and/or drug problems in all patients
4.3 identify patient factors associated with increased risk for suicide or suicide attempt
4.4 assess stressors for the patient that may have caused the attempt

handle patients who were not cooperative or refused to be assessed
re-assess patients who were not alert at the time of the initial evaluation
assess other people significant to the patient
provide treatment and aftercare
provide information to aftercare therapists

Each criterion was scored positive if the guideline gave any instructions relevant to the specific issue.

B. Quality of local guidelines

To evaluate the methodological quality of the local guidelines, the Appraisal of Guidelines for Research and Education (AGREE) instrument was used (2001). This validated instrument has been developed by an international group of guideline experts and consists of 23 key items organized in six domains. For most domains, Cronbach’s α varied between 0.64 and 0.88 (2003). It has been used in studies to evaluate the quality of clinical practice guidelines for lung cancer diagnosis and treatment (Harpole et al., 2003), guidelines for the management of major depressive disorder in the general hospital (Voel linger et al., 2003), as well as European psychiatric treatment guidelines (Stiegler, Rummel, Wahlbeck, Kissling, & Leucht, 2005). Domains and items are as follows:

1. Scope and purpose (3 items). This domain scores the presence of specific descriptions of the overall objectives, the clinical questions covered, and the patients for whom the guideline is meant to apply.

2. Stakeholder involvement (4 items). This domain scores whether all relevant professionals participated in developing the guideline, whether the patient’s views and preferences were sought after, whether the target users were defined, and whether the guideline was pilot tested among users.

3. Rigor of development (7 items). This domain scores whether systematic methods were used to search for evidence; whether the criteria for selecting the evidence and the methods used to formulate the recommendations were
clearly described; whether an explicit link was made between the recommendations and the supporting evidence; whether benefits, side effects, and risks were considered when formulating the recommendations; whether the guideline was externally reviewed by experts prior to publication; and whether a procedure was provided for updating the guideline.

4 Clarity and presentation (4 items). This domain scored whether the recommendations were specific and unambiguous, whether the different management options were clearly presented, whether key recommendations were easily identifiable, and whether the guideline was supported with tools for application.

5 Applicability (3 items). Issues pertinent to guideline implementation were evaluated in this domain. Specific factors included organizational barriers, cost implications, and monitoring criteria.

6 Editorial independence (2 items). This domain scored whether conflicts of interest were recorded and whether the guideline was editorially independent. This domain was not used in this study because it was considered irrelevant considering the subject. Some guidelines stated that the hospital administration or the medical staff had mandated that a group of cooperating professionals such as psychiatrists, nursing personnel, and managers develop the guideline.

The scores for each domain were obtained by summing up all the scores on an individual item in a domain and then standardizing them as follows:

\[
\text{obtained score} = \text{minimum possible score} \times 100\% \\
\text{maximum possible score} - \text{minimum possible score}
\]

The maximum possible score for each domain was the number of questions multiplied by the number of reviewers multiplied by four (i.e., the score for ‘strongly agree’). The minimum possible score for a domain was the number of questions multiplied by the number of reviewers multiplied by one (i.e., the score for ‘strongly disagree’).

The final component of the AGREE instrument involves making a recommendation regarding the use of the guidelines in practice. The four categories are strongly recommended, recommended (with provisos or alterations), would not recommend, or unsure.

Three reviewers (B.V., J.v.W., and G.G.) independently scored the AGREE instrument to evaluate the quality of the local guidelines. \( \kappa \) statistics were calculated for the agreement on recommendations of the guidelines, and the
intraclass correlation coefficients were calculated for absolute agreement on the five domain scores. We used a mixed-effects model, because the only raters of interest were the three that participated in the study.

As in the Netherlands university hospitals in general are more committed to the development of guidelines, a distinction between university and general hospitals was made.

RESULTS

The overall response rate to the questionnaire was 90.7%. All eight university hospitals and 80 of the 89 general hospitals responded.

Availability of local guidelines, dates, updating, and evaluation of observance
Five out of eight university hospitals reported they used local guidelines, but only four guidelines (50%) were submitted for examination. One guideline was more than ten years old; the three others were not older than five years. Two had never been updated. Two had been updated within the previous five years. One university hospital reported that they regularly evaluated staff compliance with their guideline.

Twenty-nine of the 80 (36.3%) responding general hospitals reported using local guidelines; 23 (28.8%) submitted them for evaluation. Seventeen (73.9%) guidelines were more than five years old. Eight (34.8%) had been updated within the previous five years. Nine general hospitals (39.1%) stated they regularly evaluated staff compliance with their local guidelines.

Significant differences between the university and general hospitals were found only for the criterion ‘re-assessing non-alert patient’ (χ² test: means 50% and 4.3%, respectively; P = 0.01) and the AGREE domain ‘clarity and presentation’ (t-test: means 50.7% and 67.4%, respectively; P = 0.05). However, the small number of university hospitals that submitted guidelines and the marginal differences observed between the two types of hospitals prevent meaningful comparisons. Therefore, we present the sum of the results. In total, 34 out of 88 (38.6%) hospitals reported using guidelines, and 27 of the guidelines were available for this study.

Criteria related to the content of the local guidelines (Table 4)
In 13 out of 27 (48.1%) guidelines, recommendations were made to guarantee the patient’s safety. In almost half of the local guidelines, instructions were given on how to respond to the patient. In 15 out of 27 (55.6%) guidelines, the
Table 4. Criteria for the assessment of suicide attempters in Dutch university and general hospitals

<table>
<thead>
<tr>
<th>Criterion</th>
<th>University hospital n = 4</th>
<th>General hospital n = 23</th>
<th>All n = 27</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>Addressing safety</td>
<td>3 (75)</td>
<td>10 (43.5)</td>
<td>13 (48.1)</td>
</tr>
<tr>
<td>Response to patient</td>
<td>3 (75)</td>
<td>10 (43.5)</td>
<td>13 (48.1)</td>
</tr>
<tr>
<td>Somatic consultation</td>
<td>3 (75)</td>
<td>13 (56.5)</td>
<td>16 (59.3)</td>
</tr>
<tr>
<td>Psychiatric consultation</td>
<td>3 (75)</td>
<td>18 (78.3)</td>
<td>21 (77.8)</td>
</tr>
<tr>
<td>Assessment of suicidality</td>
<td>2 (50)</td>
<td>10 (43.5)</td>
<td>12 (44.4)</td>
</tr>
<tr>
<td>Diagnosing psychiatric disorder</td>
<td>2 (50)</td>
<td>8 (34.8)</td>
<td>10 (37.0)</td>
</tr>
<tr>
<td>Detecting risk factors</td>
<td>2 (50)</td>
<td>9 (39.1)</td>
<td>11 (40.7)</td>
</tr>
<tr>
<td>Detecting stressors</td>
<td>2 (50)</td>
<td>8 (34.8)</td>
<td>10 (37.0)</td>
</tr>
<tr>
<td>Handling refusing patient</td>
<td>2 (50)</td>
<td>15 (65.2)</td>
<td>17 (63.0)</td>
</tr>
<tr>
<td>Assessing significant others</td>
<td>2 (50)</td>
<td>13 (56.5)</td>
<td>15 (55.6)</td>
</tr>
<tr>
<td>Re-assessment non-alert patient*</td>
<td>2 (50)</td>
<td>1 (4.3)</td>
<td>3 (11.1)</td>
</tr>
<tr>
<td>Regulations for aftercare</td>
<td>2 (50)</td>
<td>11 (47.8)</td>
<td>13 (48.1)</td>
</tr>
<tr>
<td>Reportage</td>
<td>2 (50)</td>
<td>14 (60.9)</td>
<td>16 (59.3)</td>
</tr>
</tbody>
</table>

* p = 0.01 ($\chi^2$-test) (Difference between university and general hospitals)

recommendation was provided that all patients should be examined by a somatic specialist, and psychiatric consultation was instructed in 21 out of 27 (77.8%). The necessity of assessing current suicidality was published in 12 out of 27 (44.4%) guidelines. In 10 out of 27 (37.0%), a psychiatric examination was recommended to detect mental illness, alcohol abuse, and drug problems. The need to detect factors associated with increased risk for suicide or suicide attempt was mentioned in 11 out of 27 (40.7%) guidelines. Identifying stressors for the patient that gave rise to the suicide attempt was called for in 10 out of 27 (37.0%).

Instructions on how to manage uncooperative or refusing patients were provided in 17 out of 27 (63.0%) guidelines. The instruction to re-assess patients who were not alert at first consultation was found in 3 out of 27 (11.1%) guidelines. Assessing significant others was instructed in 15 out of 27 (55.6%). In 13 out of 27 (48.1%) guidelines, recommendations were given for quick referral to aftercare providers. In 16 out of 27 (59.3%), procedures for handing information over to these providers were given.
Table 5  Intraclass correlation coefficients of the raters

<table>
<thead>
<tr>
<th>Domain</th>
<th>ICC</th>
<th>LB</th>
<th>UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope and purpose</td>
<td>0.73</td>
<td>0.56</td>
<td>0.86</td>
</tr>
<tr>
<td>Stakeholder involvement</td>
<td>0.52</td>
<td>0.29</td>
<td>0.72</td>
</tr>
<tr>
<td>Methodology</td>
<td>0.66</td>
<td>0.46</td>
<td>0.81</td>
</tr>
<tr>
<td>Clarity and presentation</td>
<td>0.36</td>
<td>0.11</td>
<td>0.60</td>
</tr>
<tr>
<td>Applicability</td>
<td>0.32</td>
<td>0.09</td>
<td>0.56</td>
</tr>
</tbody>
</table>

**ICC** = intraclass correlation coefficient

**LB** = lower bound; **UB** = upper bound

---

Figure 1  Percentage of guidelines for the assessment of suicide attempters scoring low (0-30%), medium (30-60%) and high (60-100%) in five out of six domains of the AGREE instrument

---

CHAP TER 3  41
Ratings in five of six domains of the AGREE instrument (Figure 1).
For the AGREE domain ‘scope and purpose’, the mean score was 43.3 (SD 29.2) with 9 out of 27 local guidelines scoring > 60%. The mean score for the domain ‘stakeholder involvement’ was 22.4 (SD 17.5) with only one guideline scoring > 60%. None of the guidelines was pilot-tested among users. The mean score for ‘rigor of development’ was 11.8 (SD 11.0) with two guidelines scoring between 30% and 60%. None of the guidelines indicated that a systematic literature study had been out of performed during its development. The mean score for the domain ‘clarity and presentation’ was 64.9 (SD 16.0) with 18 out of 27 guidelines scoring > 60%. For the domain ‘applicability’ the mean score was 14.9 (SD 12.3) with only two guidelines scoring > 30% and none more than 60%.
Table 5 shows the intraclass correlation coefficients of the three raters (two psychiatrists, B.V. and J.v.W., and a quality assurance officer, G.G.). The best agreement was for ‘scope and purpose’. A reasonable agreement was also present for ‘rigor of development’ and for ‘stakeholder involvement’. Agreement was poor on ‘clarity and presentation’ and on ‘applicability’. There was better agreement between the psychiatrists for the domains ‘scope and purpose’, ‘rigor of development’, and ‘stakeholder involvement’ than between either psychiatrist and the quality assurance officer. For these three items, the correlation coefficient was over 0.85 between the psychiatrists. However, the agreement was poor among all three for the domains ‘clarity and presentation’ and ‘applicability’, with correlations between two raters ranging from 0.30 to 0.43 for ‘applicability’ and from 0.38 to 0.58 for ‘clarity and presentation’.

Overall assessment
The reviewers agreed that 10 local guidelines (37.0%) should be recommended with provisos or alterations (one university and ten general hospital guidelines) and one should be strongly recommended. The agreement of the overall assessment was substantial (κ = 0.72).

Discussion
This report describes the first systematic study on the availability, content, and quality of guidelines for the assessment of suicide attempters in university and general hospitals in the Netherlands. The response to a written request was remarkably high. Only a minority (38.6%) of the hospitals reported using local guidelines for the assessment of suicide attempters. This result is remarkable because a national guideline was published in the Netherlands in 1991. Only one hospital reported using this national guideline, and some others men-
tioned using it to develop their own. These results are even lower than the 60% of trusts in England that had a clear policy document for dealing with deliberate self-harm (Slinn et al., 2001). We are not aware of studies in the USA addressing this subject.

Considering that half of all guidelines are outdated after 5.8 years (Shekelle et al., 2001), many of the examined local guidelines were antiquated. Moreover, most were not updated and staff compliance was seldom evaluated.

Local guidelines varied widely in their primary objectives, general length, narrative form and readability. Some addressed mainly patient safety and physician responsibilities, while others focused more on suicidality assessment. Remarkably, the necessity for psychiatric consultation for every suicide attempter, as directed by major professional groups (American Psychiatric Association, 2003; Centraal Begeleidingsinstituut voor de Intercollegiale Toetsing, 1991; Royal College of Psychiatrists, 2004), was not recommended in all local guidelines. The specific tasks of the psychiatrist were described in less than 50% of the guidelines. Other criteria for assessment, as defined in this study, were found in half of the guidelines or less than half, perhaps because most guidelines were outdated and our evaluation criteria were developed from the recent literature. On the other hand, most of the defined criteria used in this study were already in the Dutch national guideline from 1991. So most of the guidelines were incomplete even according to the 1991 national criteria. Based on analysis of the AGREE scores, the reviewers agreed to recommend (with provisos or alterations) eleven local guidelines. Taking into account the very low scores in three out of five domains (‘stakeholder involvement’, ‘rigor of development’, and ‘applicability’), this is a remarkable result. Presumably the higher scores in the domains ‘scope and purpose’ and ‘clarity and presentation’ resulted in an overall mild conclusion. The very low scores in the domain ‘rigor of development’ are the result of the fact that hardly any guideline referenced published international guidelines or other literature. If references were given, no description of the search and selection process was included. Descriptions of the methods used to formulate the guidelines were not found, and the possible effects of the guideline were seldom assessed. The low scores for the domain ‘stakeholder involvement’ probably mirror the poor presentation of the guidelines, rather than the real absence of professionals in the process of guideline development. Patient views were not mentioned in any of the guidelines.

To establish our results, we compared them with prior studies on quality of guidelines using the AGREE instrument. In a study reviewing 51 clinical practice guidelines for lung cancer diagnosis and treatment (Harpole et al., 2003), only 19 (37%) were recommended, which is comparable with the per-
centage of guidelines for assessment of suicide attempters that were recommended in our study. However, in the lung cancer guidelines the mean scores in 4 out of 5 domains were substantially higher. In a study on 61 European psychiatric treatment guidelines (Stieglert et al., 2005), using the AGREE instrument, mean scores in 4 out of 5 domains were also substantially higher. The sum of the mean scores of the domains in the lung cancer guidelines and the European psychiatric treatment guidelines were respectively 16 and 10% higher.

Many of the local guidelines failed to address issues of implementation and monitoring. Addressing such issues is necessary to take the guidelines into practice successfully. For example, an intervention study showed that teaching emergency department staff how to handle suicide attempters led to improvements in the quality of the assessment of these patients (Crawford, Turnbull, & Wessely, 1998).

The intraclass correlation coefficients between raters were very poor for some domains, which may be due to differences in interpretation of several items where the instructions were broad. Some items were relatively straightforward, as in ‘rigor of development’ and the intraclass correlation coefficient was substantial at that point (0.66). The raters agreed substantially that one guideline should be strongly recommended, and ten guidelines should be recommended (with provisos and alterations) ($\kappa = 0.72$). According to the instructions for the AGREE instrument, the overall assessment of guidelines can be the result of personal judgment or by calculating the scores in the domains. When in the majority of the domains the scores are above 60%, a guideline can be strongly recommended, and when scores are between 30-60% a guideline can be recommended (with provisos and alterations). Using this method, none of the evaluated guidelines could be strongly recommended and twelve could be recommended (with provisos and alterations). The same eight local guidelines were recommended using both methods.

In the AGREE domain ‘clarity and presentation’, among other elements ‘the precise and concrete description of which management is appropriate in what situation and in what patient group’, and ‘the different possible options for screening, prevention, diagnosing and treatment’ are appraised; in this study a mean score of 64.9% was found. However, content of guidelines was also evaluated by defined criteria and most of them were found in less than 50% of the guidelines. So only appraising content with the scores in the AGREE domain ‘clarity and presentation’ in this study had given a very positive result.

Although the majority of these Dutch hospitals did not have local guidelines, we cannot conclude from this descriptive study that physicians working in hospitals without guidelines are not capable of assessing suicide attempters
properly. And although guidelines can be a valuable tool to improve quality of care, their development does not ensure their use in practice (Feder, Eccles, Grol, Griffiths, & Grimshaw, 1999). In fact, in addition to developing and implementing guidelines, other approaches are necessary as well in order to achieve improvement of quality of care in daily practice (Grol & Grimshaw, 2003). Even in the presence of national guidelines, the small number of local guidelines found in this study, as well as the incomplete content and unsatisfactory quality of these guidelines, underline that the development and implementation of guidelines for the assessment and treatment of suicide attempters requires much more attention from professionals and local and national health care organizations. In the UK only 60% of the hospitals have implemented local guidelines (Slinn et al., 2001) and in the US this is not studied yet. The recommendations for the assessment of suicide attempters directed in the published guidelines of the Netherlands, UK and US are comparable. Therefore, the findings of this Dutch study lead to questions on the availability, content and quality of local guidelines in the UK and the US.

In conclusion, local guidelines are available in a minority of the university and general hospitals in the Netherlands. If available, their content and quality are mostly unsatisfactory. Taking into account the severity of the health-care problem of suicide attempt and suicide, the development and implementation of local guidelines, even given the availability of national guidelines, should be urgently pursued.
REFERENCES


