Introduction

It is generally admitted that the prevalence of undernutrition at home is not high. However, as showed by a number of studies carried out in Europe and USA (1), it can reach half of the patients when admitted in hospital. Moreover, in a geriatric population, undernutrition is not only linked to a morbidity and mortality increase (2) but also to a bad quality of life (3). The frequency of proteino-energetic undernutrition is difficult to evaluate because often underestimated, especially in the elderly hospitalized population (4). On the other side, there is not much of unanimity concerning the tools for such an evaluation (1).

As a matter of fact, in spite of numerous instruments for undernutrition detection, none of them has shown sufficient characteristics to be used in disease detection (5). As a consequence, due to the lack of internationally recognized evaluation scale, like for the Mini Mental State which evaluates the cognitive functions (6) and for the Geriatric Depression Scale which evaluates the depression (7), Guigoz, Vellas and Garry (8) have developed and validated an evaluation scale for the nutritional state of the elderly entitled "Mini Nutritional Assessment" (MNA), completing in that way one of the aspects of gerontologic assessment.

According to these authors, the MNA conception has to found upon the following criteria : good reproducibility, clear definition of "good" and "bad" nutritional state, consistency with requirements of clinical practice, minimization of error risk in data recording, good acceptability by patients, and finally cheapness and rapidity.

As described in a literature review by Reuben (5), the efficiency of such a tool depends on its reliability when used by different observers. The aim of the present study is to measure the interobserver agreement on MNA nutritional scale for hospitalized elderly patients. Such an assessment does not seem to have so far been carried out, but it is of major importance if MNA is wanted to be included in systematic examination of elderly people when admitted in hospital.

Material and methods

The present investigation is part of a larger study concerning 175 patients aged of 79.7 (± 8.45) (62 men and 113 women) admitted over the period december 1994-february 1995 for acute medical pathology in two geriatric units of "Centre Hospitalier Regional" of Liège (Belgium). For each patient, in addition to usual data registered when admitted (administrative and social data, personal background), clinical examination, MNA...
nutrition assessment, KATZ scale dependance and biological analysis were available. MNA form (figure 1 in annexe) is dated and completed by personal data (name, surname, age and sex).

Afterwards, the value of the answer to each item is indicated in the corresponding space. The total of points obtained by MNA goes from 0 to 30. According to the obtained score, the patient is classed in one of three types of nutritional state: satisfactory nutritional state (MNA > 24), undernutrition risk (MNA : 17 to 23.5) and bad nutritional state (MNA < 17).

The evaluation of interobserver agreement has been made with the first 20 and the last 20 patients of the general study, enabling in that way to eventually show a possible variation of the agreement level with time. For each of those patients, the MNA has been independently performed by two geriatry clinicians in charge of the two hospitalization rooms. A patient of the second class had to be discharged because it was impossible to complete the MNA by one of the two observers, due to lack of a family help.

The material of this investigation consists of 39 patients aged 79 (± 9), 12 males and 27 females. Among them, 54% come from their home (alone or with near relative), 23% live in a family and 23% stay in home for elderly or in home for elderly where medical care is provided.

The MNA assessment is performed during the first week of hospitalization. The patients considered in this study have in any case neither been transferred from another ward or from another hospital, nor spent more than 24 hours in a emergency unit.

Concerning the MNA form (figure 1), each of the 39 patients is questioned by both observers. The weight (kg) is measured by hospital attendants during the first 24 hours, the patient being undressed and with empty stomach. For all patients, the height (cm) is calculated from the age (years) and the knee height (KH, cm) according to Chumlea's formula (9):

\[
\text{Men: } 64.19 - (0.04 \times \text{age}) + (2.02 \times \text{KH})
\]
\[
\text{Women: } 84.88 - (0.24 \times \text{age}) + (1.83 \times \text{KH})
\]

In conditions where the patient is lying on his or her back, left knee up with a 90° angle between leg and thigh, the KH is measured between the heel and the femoral condyles. The brachial circumference (BC, cm) is measured in the middle of the forearm, elbow at right angle, i.e. halfway between the top of olecranon and acromion. The calf circumference (CC, cm) is measured at the largest perimeter by moving a flexible tape-measure.

All the data are gathered from the body left side, using a flexible tape-measure and only one time.

In case where it is difficult to interrogate the patient (madness, language problem, ...) the MNA is completed with the help of a relative. In such a case, the subjective evaluation of nutrition and health (items 17 and 18) is arbitrarily noted respectively 1 and 0.5 point.

### Statistical analysis

In order to estimate the agreement level between both clinicians, we have calculated the Kappa index proposed by Cohen (10) and given by the relation:

\[
K = \frac{Po - Pe}{1 - Pe}
\]

where Po is the observed proportion of agreements and Pe the proportion of agreements due to chance. A K = 0 value means an agreement only attributable to chance and K = 1 in case of perfect agreement. As a result, agreement between both observers is better and better in so far as K is close to 1.

The estimated K index can be associated to its error type s (K) which enables to check if K is statistically different from zero. For that purpose, the ratio K/s(K) is calculated and compared to the critical threshold of the normal law. Similarly two Kappa coefficients can be compared using a Chi-square test.

The method not only enables to calculate a total Kappa coefficient but also a Kappa coefficient for each MNA item, and even for each heading. This allows particularly to see if two headings close to each other for a same item can really be distinguished or if they can be considered together.

For the continuous variables (height, brachial circumference, calf circumference) and also for the MNA score (0-30), the determination of the interobserver agreement is obtained by doing a two-criteria variance analysis (ANOVA-2) and calculating the "interclass correlation coefficient" (11). The value of the latter is always between 0 and 1, and the closer to 1, the better the agreement between both experimenters. We give these values just for information.

All statistical tests are considered as significant for an uncertainty level of 5% (p < 0.05).

### Results

#### Concordance degree of total MNA score

The Cohen's Kappa coefficient for the total MNA score is 0.587 (± 0.162) for the first period of this investigation and 0.412 (± 0.166) for the second one. A statistical test shows that there is no significant difference between these two values. Consequently the whole analysis has been
performed on the group of 39 patients considered in only one single group.

The distribution per classes of MNA scores obtained by both clinicians is given in Table 1.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution per classes of MNA scores obtained by both clinicians for the 39 examined patients</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>M N A</th>
<th>CLINICIAN 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 17</td>
<td>5</td>
</tr>
<tr>
<td>17-23,5</td>
<td>3</td>
</tr>
<tr>
<td>&gt; 24</td>
<td>0</td>
</tr>
<tr>
<td>CLINICIAN 2</td>
<td></td>
</tr>
<tr>
<td>&lt; 17</td>
<td>5</td>
</tr>
<tr>
<td>17-23,5</td>
<td>17</td>
</tr>
<tr>
<td>&gt; 24</td>
<td>6</td>
</tr>
</tbody>
</table>

The total Kappa coefficient obtained reaches 0.51 (95\% confidence interval : 0.28 - 0.74). Although this value is statistically different from zero. It cannot be claimed that the agreement between both interrogators be excellent, but at most satisfactory.

The Kappa calculation for each of the 3 MNA classes leads to the following results : 0.65 (bad nutritional state), 0.42 (undernutrition risk) and 0.51 (satisfactory nutritional state). As a result of this, it is observed that the agreement between both clinicians is better for the first and last classes, which is a common observation with score scales.

It is important to add that, among the 39 aged patients evaluated, none of them is found in opposite classes (1 and 3, or 3 and 1) by both clinicians (see table 1).

Concordance degree of different items of MNA

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agreement degree for the 18 items of MNA scale</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item</th>
<th>Kappa total</th>
<th>Kappa per class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1. B.M.I (1)</td>
<td>0.84</td>
<td>1.00</td>
</tr>
<tr>
<td>2. B.C. (2)</td>
<td>0.30</td>
<td>(- 0.026)</td>
</tr>
<tr>
<td>3. C.C. (3)</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>4. Weight loss</td>
<td>0.54</td>
<td>0.84</td>
</tr>
<tr>
<td>5. Indep. at home</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>6. More than 3 medicines</td>
<td>0.54</td>
<td>0.54</td>
</tr>
<tr>
<td>7. Psychological stress</td>
<td>0.57</td>
<td>0.57</td>
</tr>
<tr>
<td>8. Motricity</td>
<td>0.66</td>
<td>0.68</td>
</tr>
<tr>
<td>9. Madness/depression</td>
<td>0.44</td>
<td>0.44</td>
</tr>
<tr>
<td>10. Bed sores</td>
<td>(0.07)</td>
<td>(- 0.07)</td>
</tr>
<tr>
<td>11. Number of meals (daily)</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>12. Proteinic score</td>
<td>0.46</td>
<td>0.47</td>
</tr>
<tr>
<td>milk</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>eggs</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>13. Fruits/vegetables</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>14. Appetite loss</td>
<td>0.32</td>
<td>(0.12)</td>
</tr>
<tr>
<td>15. Hydratation</td>
<td>0.48</td>
<td>(0.04)</td>
</tr>
<tr>
<td>16. Nutritional dependance</td>
<td>0.38</td>
<td>0.79</td>
</tr>
<tr>
<td>17. Nutritional subj. evaluation</td>
<td>0.63</td>
<td>0.79</td>
</tr>
<tr>
<td>18. Health subj. evaluation</td>
<td>0.33</td>
<td>(0.26)</td>
</tr>
</tbody>
</table>

1) Body mass index, (2) brachial circumference, (3) calf circumference. Values in brackets are not significant.

The Cohen's Kappa coefficient has been calculated for each MNA item, at the total level of the item as well as for each of its headings. All the results are given in table 2.

For anthropometric indices, it is seen from table 2 that an excellent agreement is obtained for body mass index (BMI) (0.84) and calf circumference (0.72). Contrariwise, an agreement problem shows up for brachial circumference (0.30). The agreement for weight loss is good (0.54) and excellent for the heading "weight loss of more than 3 kilos" (0.84).

For the global assessment, there is a good agreement for all items, i.e. independance at home (0.7), number of medicine (0.54), stress (0.57), motricity (0.66) and psychoneurological problems (0.44), except for the item concerning bed sores where the agreement is not significant.

As for dietetic indices, an excellent agreement is obtained for the number of daily meals (1.0) and for the fruit and vegetable intake (0.74). It is satisfactory for the items : total proteinic score (0.46), appetite loss (0.32) and nutritional dependance (0.38). On the contrary, no agreement is observed for the number of beverage glasses taken daily (0.12), except for the heading "less than 3 glasses per day" where a satisfactory agreement is to be noted (0.48).

Concerning the health subjective evaluation, the agreement is better for the item referring to nutrition (0.63) than for the one referring to health (0.33).
Discussion

Setting up a tool such as MNA scale, in order to evaluate the nutrition state of the elderly has its grounds, among other things, on reproducibility as well as on answers to a list of questions which do not depend on the operator. To our knowledge, no previous study has been so far carried out to confirm this.

In the present investigation, we showed that, if results recorded by both experimenters are compared, the agreement level (calculated by Cohen's Kappa index) between MNA total scores is statistically significant but middle (Kappa = 0.51). This unperfect interobserver agreement can ben explained by the following reasons.

Firstly, the answer of an old person to a precise question may vary with time, as has been criticized by Samet (12) in nutritional evaluation requesting short-time memory, especially if he or she presents a major or minor communication trouble (deafness, problem due to different language, ...). In the present case, the answer variability is perhaps worsen by the effect produced by acute pathology for which the old person is admitted in a hospitalization unit (diabetes, bronchopneumonia, ...).

Secondly, a better knowledge of the clinical background of the patient by the observer who took him or her in charge could influence the answer to certains items. For example, one of the questioned patients disclaimed his illness, influencing in that way the answer to the items, the resulting total MNA score, and also possibly the type of class to wich he belonged.

The total MNA score enables one to define three classes of elderly : satisfactory nutritional state, undernutrition risk and bad nutritional state. The comparative study of the results obtained for the three classes has enabled to demonstrate a better agreement degree, as indicated by a larger Kappa value, for the first and last classes : 0.63 for the bad nutritional state and 0.51 for the satisfactory nutritional state. The middle class (undernutrition risk) is consequently more liable to errors of judgment.

In an attempt to understand these differences, it seemed interesting to us to make a finer analysis of items for which the Kappa value is not significant.

Critical analysis of questionable items

The first item for which a disagreement is noted is the brachial circumference (item 2) measure with a very low Kappa value (0.30) whereas that of the calf circumference is definitely significant (0.72). Such a difference is surprising but the difference between both these results is probably linked on one side to the arm texture, the elasticity of which varies with age and to the difficulty to localise, very accurately, the anatomic guide marks (13).

However, Chumlea has demonstrated the reproducibility of the brachial circumference measure (14), this being made three times and the three measurements average, which is not taken into account in our methodology. On the other side, whereas the values of the exact measures of brachial and calf circumferences seem to agree, the agreement degree of their scores is very different. This comes from the fact that the brachial circumference is too strict, i.e. for only one cm gap, class 1 changes to class 3.

The second item for which the agreement degree is not significant is the presence of bed sores or skin sores (item 10). An error of one point is easily made because the answer depends on the feeling of the sore intensity in such a way that the observer may doubt there is really one.

Finally, the Kappa value for the item "hydratation" (item 15) is not significant. Nevertheless, a very good agreement (0.48) for the answer "less than 3 glasses per day" (bad hydratation) is observed, whereas the headings "3 to 5 glasses" and "more than 5 glasses" remain difficulty distinguishable.

In the same way, items 4 (weight loss), 9 (neuro psychological problems), 14 (anorexia), 16 (nutritional self-governing) and 18 (health subjective estimation) have altogether a kappa value statistically significant but low.

The reason of this probably lies in the difficulty either to distinguish headings (example : weight loss = 0 or between 1 and 3 kilos), either to choose the strictness degree attributed to certain headings without objective measure (example : madness/mild or hard depression). Moreover, concerning the weight loss, Morgan's study (15) has clearly demonstrated that, for the estimation of their weight loss, 25 % of the patients thought they had some although it was not the case.

As a conclusion of the above discussion, items 2, 12 and 15 can lead to an error of 0.5 point and items 9, 10, 14, 16 and 18 to an error of 1 point. On the whole, the points difference between observers can therefore easily reach 6.5 points. However, even in spite of such a disagreement, classes differing by 8 points are being well distinguished.

Conclusion

It is important that the clinician be aware of the validity limits of his tool. The present study demonstrates that the evaluation of the nutritional scale by MNA is reliable, even if a number of items rise question and could be improved.

Its cheapness as well as the performance speed make it an interesting tool in geriatriy practice. However, there are some limits. The attention of theclinician has to be drawn when a patient has to be considered in the intermediate class (undernutrition risk) because of its worse characterisation. If such is the case, a further investigation could be necessary, either by another nutritional scale, either also by a deeper nutritional anamnesis or finally by biological tests.

In a similar way, difficulties have been met in nutritional evaluation in cases where the old person is mad or without close family. As a consequence, in such circumstances, the present tool should be carefully used.

In addition to the results presented in this study, a further analysis should be carried out in the future in order to check the effect, on the total MNA score, of knowing the medical background and the occupation.

References


### Mini Nutritional Assessment (MNA™)

**MINI NUTRITIONAL ASSESSMENT MNA™**

<table>
<thead>
<tr>
<th>ID#</th>
<th>Last Name</th>
<th>First Name</th>
<th>M.</th>
<th>Sex</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Weight, kg</td>
<td>Height, cm</td>
<td>Knee Height, cm</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ANTHROPOMETRIC ASSESSMENT**

1. Body Mass Index (BMI) (weight in kg) / (height in m)²
   - a. BMI = 19 = 0 points
   - b. BMI 19 to < 21 = 1 point
   - c. BMI 21 to < 23 = 2 points
   - d. BMI ≥ 23 = 3 points

2. Mid-arm circumference (MAC) in cm
   - a. MAC < 21 = 0 points
   - b. MAC 21 to < 22.5 = 0.5 points
   - c. MAC ≥ 22.5 = 1.0 points

3. Gait circumference (GC) in cm
   - a. GC < 31 = 0 points
   - b. GC > 31 = 1 point

4. Weight loss during last 3 months
   - a. weight loss greater than 5 kg (5.6 lbs) = 0 points
   - b. does not know = 1 point
   - c. weight loss between 1 and 3 kg (2.2 to 6.6 lbs) = 2 points
   - d. no weight loss = 3 points

**GENERAL ASSESSMENT**

5. Lives independently (not in a nursing home or hospital)
   - a. no = 0 points
   - b. yes = 1 point

6. Takes more than 3 prescription drugs per day
   - a. yes = 0 points
   - b. no = 1 point

7. Has suffered psychological stress or acute disease in the past 3 months
   - a. yes = 0 points
   - b. no = 2 points

8. Mobility
   - a. bed or chair bound = 0 points
   - b. able to get out of bed/chair but does not go out = 1 point
   - c. unable to get out = 2 points

9. Neurological problems
   - a. severe dementia or depression = 0 points
   - b. mild dementia = 1 point
   - c. no psychological problems = 2 points

10. Pressure sores or skin ulcers
    - a. yes = 0 points
    - b. no = 1 point

**DIETARY ASSESSMENT**

11. How many full meals does the patient eat daily?
    - a. 1 meal = 0 points
    - b. 2 meals = 1 point
    - c. 3 meals = 2 points

**POINTS**

12. Selected consumption markers for protein intake
    - a. At least one serving of dairy products (milk, cheese, yogurt) per day? yes no
    - b. Two or more servings of legumes or eggs per week? yes no
    - c. Meat, fish or poultry every day? yes no
    - d. If 0 or 1 yes = 0.0 points
    - e. If 2 or 3 yes = 0.5 points
    - f. If 3 yes = 1.0 points

13. Consumes two or more servings of fruits or vegetables per day?
    - a. no = 0 points
    - b. yes = 1 point

14. Has food intake declined over the past three months due to loss of appetite, digestive problems, chewing or swallowing difficulties?
    - a. severe loss of appetite = 0 points
    - b. moderate loss of appetite = 1 point
    - c. no loss of appetite = 2 points

15. How much fluid (water, juice, coffee, tea, milk, ...)
    - a. less than 3 cups = 0 points
    - b. 3 to 5 cups = 0.5 points
    - c. more than 5 cups = 1.0 points

16. Mode of feeding
    - a. Unable to eat without assistance = 0 points
    - b. self-fed with some difficulty = 1 point
    - c. self-fed without any problem = 2 points

**SELF ASSESSMENT**

17. Do they view themselves as having nutritional problems?
    - a. major malnutrition = 0 points
    - b. does not know or moderate malnutrition = 1 point
    - c. no nutritional problem = 2 points

18. In comparison with other people of the same age, how do they rate their health status?
    - a. not as good = 0.0 points
    - b. does not know = 0.5 points
    - c. as good = 1.0 points
    - d. better = 2.0 points

**ASSESSMENT TOTAL (max. 30 points):**

<table>
<thead>
<tr>
<th>Points</th>
<th>Malnutrition Indicator Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>= 28</td>
<td>well-nourished</td>
</tr>
<tr>
<td>17 to 23.5</td>
<td>at risk of malnutrition</td>
</tr>
<tr>
<td>&lt; 17</td>
<td>malnourished</td>
</tr>
</tbody>
</table>