Comprehensive Geriatric Assessment (CGA) and the MNA: An Overview of CGA, Nutritional Assessment, and Development of a Shortened Version of the MNA

Laurence Z. Rubenstein, Judith Harker, Yves Guigoz, and Bruno Vellas

GRECC (Geriatric Research Education and Clinical Center), Sepulveda VA Medical Center, Los Angeles, Calif., USA; Nestlé Research Center, Nestec Ltd., Lausanne, Switzerland; Department of Internal Medicine and Clinical Gerontology, CHU Purpan-Casselardit, Toulouse, France

What Is Comprehensive Geriatric Assessment?

Comprehensive geriatric assessment (CGA) is a multidimensional interdisciplinary diagnostic process intended to determine a frail elderly person’s functional capabilities and medical and psychosocial problems, in order to develop an overall plan for treatment and long-term follow-up. While CGA includes many components of the standard medical diagnostic evaluation, it focuses on the special needs of frail elderly persons. As such, CGA goes well beyond the routine examination in its emphasis on functional status and quality of life, its thoroughness, and its use of standardized measurement instruments and interdisciplinary teams [1].

CGA usually incorporates at least four broad assessment domains: medical status, functional status, psychological function, and socioenvironmental indices. Each domain contains a number of subdomains, and most are assessed with the assistance of standardized scales and instruments outlined in Table 1. For example, psychological function contains at least the major subdomains of cognition and affect, each of which can be assessed by a number of standardized scales (for example, the mini-mental state examination for cognition and the geriatric depression scale for affect). Medical status contains many subdomains corresponding to specific organ systems (for example, cardiac, pulmonary) as well as...
### Table 1. Measurable domains of comprehensive geriatric assessment with examples of possible measurement scales*

<table>
<thead>
<tr>
<th>Domain/subdomain</th>
<th>Essential concepts</th>
<th>Examples of useful scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global functional status</td>
<td>Basic activities of daily living (ADL)</td>
<td>Self-care, basic mobility, incontinence</td>
</tr>
<tr>
<td></td>
<td>Instrumental activities of daily living (IADL)</td>
<td>Household activities, shopping, finances, etc.</td>
</tr>
<tr>
<td>Mental health</td>
<td>Cognitive status</td>
<td>Cognition, mental function, memory, etc.</td>
</tr>
<tr>
<td></td>
<td>Affective status</td>
<td>Depression, anxiety, well-being</td>
</tr>
<tr>
<td>Social activities and supports</td>
<td></td>
<td>Social network, community activities, etc.</td>
</tr>
<tr>
<td>Physical function</td>
<td>Mobility – gait and balance</td>
<td>Gait, balance, fall risk</td>
</tr>
<tr>
<td></td>
<td>Nutritional adequacy</td>
<td>Nutritional status and risk</td>
</tr>
</tbody>
</table>

* Adapted from Table 15-1 in Rubenstein LZ and Rubenstein LV [2].

Functional groupings (for example, special senses, gait and balance). An important subdomain of medical status is nutritional status, and this too contains subcomponents. These are dealt with in detail elsewhere in this volume.

CGA is performed in a variety of locations and health care contexts, usually in conjunction with patient admission or enrollment in a program and often repeated or updated at periodic intervals thereafter. Not only is it a fundamental part of care on special geriatric inpatient units, but it is a basic part of most other specialized geriatric care programs (for example, home care, day hospital, rehabilitation programs, geriatric consultation services, community preventive health programs for seniors), occurring in conjunction with patient entry and often repeated or updated at periodic intervals thereafter. It is also performed in many
primary care settings that care for older patients, as a supplement to the standard medical evaluation.

However, the process of CGA is not the same in all these settings and can vary considerably in complexity. It can be viewed as a continuum, ranging from a limited screening assessment by primary care physicians or community health workers focused on identifying an older person’s functional problems and disabilities, to more thorough evaluation of these problems by a geriatrician or interdisciplinary team coupled with initiation of a therapeutic plan in specialized centers. Because it has become so much an integral part of care in these geriatric programs, it is often difficult to distinguish the benefits of the CGA process itself from those of the treatment, rehabilitation, and follow-up processes which follow the CGA and are often provided in the same settings. Indeed, when discussing CGA and its benefits, one usually considers CGA together with its entire geriatric care and follow-up program. This broader context of CGA creates both a challenge for interpretation and an opportunity for greater program success.

The Benefits of CGA

A growing body of literature supports the effectiveness of CGA programs in several different settings. Early descriptive studies of CGA programs in hospitals and the community indicated several important benefits, such as improved diagnostic accuracy, reduced discharges to nursing homes, increased functional status, and reduced drug treatment. However, these early studies did not have concurrent controls, and thus could not distinguish the effects of the programs from the simple effects of improvement over time. Nor was it clear how these apparent benefits, most of which affected process of care, would relate to longer-term outcome benefits.

Beginning in the 1980s, controlled studies began to emerge which corroborated some of the earlier descriptive studies and documented additional benefits to both process of care and to short- and long-term outcomes. Many of these controlled trials showed that programs led to improved survival, decreased use of institutional services, and improved levels of both mental and physical functioning. However, there was considerable diversity in trial results, with many showing a high degree of program benefit, while others reported minimal effects [2, 3].

To resolve some of this discrepancy, and to try to identify particular program elements associated with particular benefits, systematic meta-analysis has been performed [4]. This meta-analysis included published data from 28 controlled trials, as well as substantial amounts of previously unpublished data systematically retrieved from many of the studies. The meta-analysis subdivided CGA programs into five types: hospital units (n = 6 trials), hospital consultation teams (n = 8), in-home assessment services (n = 7), outpatient assessment services (n = 4), and “hospital-home assessment services” (n = 3), which performed in-home
assessments on patients recently discharged from hospitals. The meta-analysis confirmed many of the major reported benefits for many of the program types. Among these statistically and clinically significant benefits included: reduced risk of mortality (by 22% for hospital-based programs at 12 months, and by 14% for all programs combined at 12 months); improved likelihood of living at home (by 47% for hospital-based programs and by 26% for all programs combined at 12 months); reduced hospital admissions (by 12% for all programs at study end); greater chance of cognitive improvement (by 47% for all programs at study end); and greater chance of physical function improvement for patients on hospital units (by 72% for hospital units). In general, hospital units and in-home assessment programs showed the strongest and greatest variety of benefits. Covariate analysis showed that programs with actual clinical control of medical recommendations (as opposed to simply making consultations) and those providing extended follow-up were more likely to be effective.

Studies published since the meta-analysis have been, in general, consistent with the earlier results. Ongoing studies are continuing to examine the benefits of CGA in different program models, attempting to define the essential elements, the most appropriate patients, and the most cost-effective combinations.

**Who Benefits Most from CGA? The Need for Patient Targeting**

Since CGA programs can be expensive interventions, and since clearly not all elderly persons can measurably benefit from CGA, it is important to identify criteria that will help identify persons most likely to be helped. These are often called “targeting criteria,” since they are intended to use clinical information to direct or “target” the most appropriate patients to a given program. Ideally, targeting criteria should separate all potential patients into at least three groups: persons appropriate for the program, those too healthy or independent, and those too dysfunctional or with too poor a prognosis. Actually, this concept of targeting is really a triage function, no different for geriatrics than selecting patients for most other clinical interventions. Most surgical procedures, for example, have a specific list of indications and contraindications, which could be used to identify persons as “too sick,” “too well,” and “appropriate.” However, the targeting process has gained more attention within geriatrics, perhaps because of the relative scarcity of geriatric resources and because geriatric services had formerly been less selective of patients.

The nature of the targeting criteria differs between programs, and should reflect program purposes, structure, and costs. For example, hospital-based CGA units aimed at specific groups of patients (for example, stroke patients, patients with hip fracture, patients expecting a long stay) need more stringent and specific criteria than do community outreach programs that provide CGA to independent and relatively healthy elderly persons. As a rule, relatively inexpensive screening-
type programs require only general and inclusive criteria, intended to admit a large proportion of elderly persons, while expensive hospital rehabilitation programs need very selective criteria.

While most targeting criteria have been derived empirically from clinical judgment, trying to select patients who would be most likely to benefit from the program, clinical trials have in some cases been able to validate the usefulness of the criteria. For example, in the Sepulveda VA randomized trial of an inpatient CGA unit [5, 6], a two phased patient identification process used selection criteria based on clinical experience. Hospital patients who remained in the hospital 7 days after admission were first screened for inclusion criteria by non-physician research assistants, looking for “persistent medical, functional, or psychosocial problems interfering with discharge home.” Patients so identified would next be screened for exclusionary criteria (for example, terminal illness, unstable medical condition) by the research assistant in consultation with a geriatric physician, who would then make the final decision. The criteria divided the screened population into several subgroups (for example, “appropriate” (about 10% of those screened), “too independent,” “terminal,” and so on), and group assignment was shown to predict major clinical outcomes (that is, death, discharge location, 1-year living location, nursing home use). Validation of these criteria was performed in several stages over subsequent years (see below).

In a Canadian trial of a hospital-based CGA consultation team, Hogan & Fox [7] subdivided all elderly hospital patients into seven categories, within 48 h of admission, based on function, mental status, and prognosis. This process identified 38% of patients for the CGA intervention (categories 3–5), while 49% were excluded as being too functional (categories 1 and 2), and 14% were excluded as being either totally dependent or terminal (categories 6 and 7). The clinical categories were shown to predict subsequent mortality, nursing home placement, and functional status measured on the Barthel scale. The CGA intervention was then tested by random assignment, and the intervention was shown to improve survival, function, and possibly utilization of services for the targeted subgroup of patients.

Some of the most precise criteria have come from retrospective subgroup analyses of controlled trials, since controlled trials can pinpoint subgroups of their study populations who were more (and less) likely to benefit from the intervention. In the Sepulveda geriatric evaluation unit (GEU) trial [8], the randomized study population was subdivided by potential baseline risk factors and examined for 1-year survival. Although virtually all subgroups showed some significant benefit or positive trends, some subgroups clearly benefited more than others. For example, patients with lower baseline scores in function and mental status showed a greater survival benefit than patients with higher baseline scores. Similarly, increased benefit was seen among patients with heart and lung diseases and those with better affect scores. To refine targeting criteria, subgroups of patients were sought who would survive only if assigned to the GEU. Two such groups
were cardiac patients with primary “geriatric” and “rehabilitation” problems, and noncardiac patients with primary “medical” problems.

In a trial of a hospital geriatric assessment unit by Applegate et al. [9], randomized patients were further stratified by their perceived risk of nursing home placement (high vs. low risk). There were greater effects from the unit on low risk stratum patients in terms of survival and functional improvement and on high risk stratum patients in terms of nursing home placement. The authors concluded that the beneficial effects of hospital CGA units on mortality and function may be greatest for elderly hospital patients with “moderate illness and moderate functional impairment.”

While benefit size and the specific subgroup criteria for patients expected to derive maximum benefit differ in the above studies, all of them clearly conclude that patients who benefit most tend to be those with some definable geriatric problem interfering with function, yet who do not have a prognosis so poor as to be essentially too late for improvement. Thus again we see the trichotomous subdivision of elderly patients.

The meta-analysis discussed earlier confirmed the importance of targeting criteria in producing beneficial outcomes in controlled trials of CGA [4]. In particular, when the use of explicit targeting criteria in a CGA trial was included as a covariate, targeting was correlated with an increase in CGA benefits among certain CGA subgroups. For example, among the hospital-based CGA studies, positive effects on physical function and likelihood of living at home at 12 months were greater in studies that excluded patients who were relatively “too healthy.” A similar effect on physical function was seen in the institutional studies that excluded persons with relatively poor prognoses. The reason for this effect of targeting on outcome size lies in the ability of careful targeting to concentrate the intervention on patients who can benefit, without diluting the effect size with persons too ill or too well to show a measurable improvement. Future research is certainly needed to refine these targeting criteria and to better tailor specific criteria to specific programs.

Assessing Nutrition in CGA: The Role of the MNA and Development of A Short Form Version, the MNA-SF

Nutritional deficiency is a common and extremely serious problem in the elderly population. While the reported prevalence of nutritional deficiency depends greatly on definitions used and populations studied, its surprisingly high rates can be considered to have reached epidemic proportions. For example, using the definition of “decreased nutrient reserves,” up to 15% of ambulatory outpatients, 35–65% of elderly hospital patients, and 25–60% of institutionalized older adults have been found to be malnourished [10, 11]. These prevalence data are described in more detail elsewhere in this volume.
It seems obvious that detection of both overt malnutrition and risk for under-nutrition must be a vital objective of CGA. The need for both screening instruments and more definitive diagnostic tools is clear, and is reflected in the proliferation of nutrition screening and assessment recommendations and tools that have emerged in the past two decades. These have been described in detail elsewhere [10, 12]. Among the most recent and extensively tested is the Mini Nutritional Assessment (MNA). The MNA fulfills many of the definitional criteria of both screening and diagnostic measures. It was developed and validated on large representative samples of elderly persons, and has been shown to be clinically useful in a variety of populations and countries for better identifying persons at nutritional risk and for providing information on dietary habits useful in intervention planning [11]. On the other hand, it is somewhat long and complex in comparison with many other screening tools, and this may impede its widespread acceptance as a short screening tool, particularly in situations where the CGA must be both brief and inclusive of multiple domains. The entire MNA was designed to take less than 20 min to administer, and was recently described as taking "about 10 min" [13], which is certainly reasonable for a diagnostic test, especially in a research or academic center. However, even 10 min might be too long to expect a primary care physician to add onto an already busy schedule for the average elderly person, in view of the many other components that typically need to be included in the comprehensive assessment – most of which have been designed to be administered in under 5 min. Moreover, several questions on the MNA require special training (for example, the anthropometry) and others are relatively subjective, requiring judgments either from the patient or caregiver, both of which might impair diagnostic accuracy. Thus a short form version of the MNA, if it could retain much of the validity and usefulness of the longer form, seems to be a worthwhile goal.

We have attempted to produce such a short form version (the MNA-SF) using a stepwise simplification process of the full MNA, trying to retain questions of the highest estimated importance, ease of data collection, face validity, and correlation both with the total MNA score and with clinical judgment of nutritional status. Our intent was to preserve as much predictive accuracy as possible while decreasing the time and training needed to administer the screen. We based our effort on the original population database of largely frail elderly persons in Toulouse, France, on which the MNA was developed and tested. Our intention is to try to subsequently validate the MNA-SF on other populations and modify it as necessary.

In attempting to select a subgroup of questions that would be retained in the MNA-SF, we used the following guidelines. We selected items that correlated well with the full MNA score and had good individual diagnostic characteristics (that is, high sensitivity, specificity, and overall diagnostic accuracy), based on independent clinical assessment of nutritional status. We chose items that, as a set, had high internal consistency. But we tried to avoid items that seemed to have
redundant content, that required special training to administer, that were particularly time-consuming, that involved difficult subjective recall, or that may have had problems with missing or “don’t know” answers. Ultimately, our aim was to identify a subscale that would approach the accuracy of the full MNA yet be brief enough for widespread use as a screening tool. One advantage of the MNA over many other nutritional assessment instruments is that it does not rely on laboratory testing, and this advantage would be retained by the MNA-SF. The ability of the MNA to help plan therapeutic interventions by defining eating habits would be lost in the MNA-SF, but if the full MNA were to supplement the MNA-SF for persons identified as being at risk, then this disadvantage could be overcome.

Methods

We used patient data from the original population in Toulouse, France, on whom the MNA was developed – the Toulouse 1991. The population has been fully described elsewhere [11], but briefly it consisted of both a hospital inpatient geriatric population (n = 105) stratified into three groups by serum albumin (<3.0 g/dl, n = 35; 3.0–3.5 g/dl, n = 35; >3.5 g/dl, n = 35), and a healthy elderly population living in the community, recruited from a community prevention program (n = 50). Characteristics of the overall group of subjects (n = 155) included the following: 66% were female; mean age was 79 years (70% ≥ 75 years); 56% had a mini-mental state score of <24/30; 43% had an ADL score of <4/6; 49% reported moderate to severe anorexia; 45% were taking more than three drugs, and 56% had serum albumin of ≤3.5 g/dl. Thus the population was largely an ill and frail hospital population combined with a smaller group of healthy elders.

Data from the full MNA on the entire population had been entered into the database, with complete MNA data available for most (n = 151) of the population. As described elsewhere [11], the MNA consists of 18 questions in four categories: anthropometric measurements (four questions), global assessment (six questions), dietary assessment (six questions), and self-perception of health and nutrition (two questions). Individual questions have weighted scores, and the scale is scored from 0 to 30. Interpretation established for the total score is: ≥24 = well-nourished, 17–23.5 = at risk of malnutrition, <17 = undernutrition.

We began the simplification process by looking for possible redundancy of items. Pearson correlation coefficients were calculated between all pairs of items. Any items with very high inter-item correlation (r > 0.80) plus overlapping item content were considered redundant. Although some pairs of items had moderately high correlations, none was redundant.

Next we looked at internal consistency, using item analysis procedures (Systat 7.0 and Crunch 4.1), with the Cronbach’s coefficient α as a measure of internal consistency. Although we recognized that nutritional status has multiple components, we assumed that these components measured aspects of a single underlying
construct, namely nutritional status. Therefore the 18 items were treated as a single scale. These procedures report the $\alpha$ for all 18 items and, for each item, the change in $\alpha$ (for the remaining 17 items) if that item were excluded. If excluding an item produced no change or an increase in $\alpha$, thereby increasing the overall internal consistency, that item was a candidate for exclusion from the item set. Successive calculations of internal consistency were done with the best remaining 15, 12, nine, and six items. As an independent check, Pearson correlations between each item and the MNA total score were examined. Items that had poor correlations to the total score were also candidates for exclusion, and the items selected were in agreement with those identified by item analysis. Finally we calculated the optimal sensitivity, specificity, and diagnostic accuracy for each item, using physician judgment of nutritional status (clinical status) as the gold standard. Items with poorest relation to clinical status were also candidates for exclusion. Again, there was good agreement with the candidates selected by item analysis.

After identifying what seemed to be the best possible short form version (MNA-SF), we tested the MNA-SF against the full MNA to see how much accuracy had been lost and to discover if there were items we might need to add back in order to improve the short form to an acceptable level. Sensitivity, specificity, and overall diagnostic accuracy for the MNA-SF were compared to those for the full MNA in ability to identify persons judged independently to have nutritional problems, as well as to identify persons with deficient serum albumin concentrations. Threshold values for the MNA-SF were chosen based on receiver-operating curves (ROCs) of diagnostic accuracy.

Results

Table 2 shows the diagnostic characteristics (sensitivity, specificity, overall predictive accuracy) of the 18 individual items of the MNA, using clinical nutritional status as the gold standard. The clinical status rating of “uncertain” ($n = 13$) was excluded from these calculations. $\chi^2$ Analysis examined the relation of each item’s response categories with clinical status (normal or malnourished). All but item No 5 were significant at $p < 0.05$. For items with the same number of response categories, the larger the $\chi^2$, the stronger the relation. For all items with more than two response categories, classification cutpoints were selected to maximize sensitivity. In some cases, two cutpoints were tried – for example, with item No 1, we tested body mass index (BMI) values of both $<21$ and $<23$ as possible indicators of malnutrition.

As can be seen, individual items were quite variable in their ability to predict overall nutritional status, with some items being quite good, that is, No 1 (BMI), No 4 (weight loss), No 7 (illness/stress), No 8 (housebound), No 9 (dementia/depression), No 14 (appetite loss), and others not very good, for example, No 2
### Table 2. Diagnostic characteristics of the individual 18 items from the MNA, relative to their ability to identify individuals with nutritional problems as determined by clinical diagnosis (as the gold standard), using the Toulouse 91 population database (n = 151)

<table>
<thead>
<tr>
<th>Item</th>
<th>Content</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>$\chi^2$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>BMI &lt;23</td>
<td>0.795</td>
<td>0.774</td>
<td>0.787</td>
<td>53.058</td>
</tr>
<tr>
<td>2</td>
<td>Mid-arm circumference ≤ 22</td>
<td>0.125</td>
<td>1.00</td>
<td>0.454</td>
<td>7.186</td>
</tr>
<tr>
<td>3</td>
<td>Calf circumference &lt; 31</td>
<td>0.614</td>
<td>0.960</td>
<td>0.745</td>
<td>45.821</td>
</tr>
<tr>
<td>4*</td>
<td>Recent weight loss &gt; 1 kg</td>
<td>0.875</td>
<td>0.774</td>
<td>0.837</td>
<td>62.690</td>
</tr>
<tr>
<td>5</td>
<td>Lives in NH or hospital</td>
<td>0.483</td>
<td>0.604</td>
<td>0.528</td>
<td>1.014 (n.s.)</td>
</tr>
<tr>
<td>6</td>
<td>Takes &gt; 3 Rx medications</td>
<td>0.584</td>
<td>0.755</td>
<td>0.648</td>
<td>15.379</td>
</tr>
<tr>
<td>7*</td>
<td>Acute illness/stress</td>
<td>0.910</td>
<td>0.906</td>
<td>0.909</td>
<td>92.556</td>
</tr>
<tr>
<td>8</td>
<td>Housebound</td>
<td>0.775</td>
<td>0.925</td>
<td>0.831</td>
<td>67.787</td>
</tr>
<tr>
<td>9</td>
<td>Dementia or depression</td>
<td>0.652</td>
<td>0.943</td>
<td>0.761</td>
<td>48.028</td>
</tr>
<tr>
<td>10</td>
<td>Pressure sores</td>
<td>0.416</td>
<td>0.981</td>
<td>0.627</td>
<td>26.694</td>
</tr>
<tr>
<td>11</td>
<td>&lt; 3 meals/day</td>
<td>0.122</td>
<td>0.962</td>
<td>0.489</td>
<td>7.346</td>
</tr>
<tr>
<td>12</td>
<td>&lt; 3 protein markers</td>
<td>0.588</td>
<td>0.736</td>
<td>0.645</td>
<td>15.174</td>
</tr>
<tr>
<td>13</td>
<td>&lt; 2 vegetables/day</td>
<td>0.165</td>
<td>1.00</td>
<td>0.486</td>
<td>9.715</td>
</tr>
<tr>
<td>14*</td>
<td>Appetite loss or eating problem</td>
<td>0.775</td>
<td>0.981</td>
<td>0.852</td>
<td>76.054</td>
</tr>
<tr>
<td>15</td>
<td>&lt; 5 cups fluid/day</td>
<td>0.565</td>
<td>0.792</td>
<td>0.652</td>
<td>17.479</td>
</tr>
<tr>
<td>16</td>
<td>Self-feeding difficulty</td>
<td>0.472</td>
<td>0.981</td>
<td>0.690</td>
<td>32.462</td>
</tr>
<tr>
<td>17</td>
<td>Self-rated malnutrition</td>
<td>0.652</td>
<td>0.962</td>
<td>0.768</td>
<td>51.331</td>
</tr>
<tr>
<td>18</td>
<td>Lower self-rated health status</td>
<td>0.618</td>
<td>0.887</td>
<td>0.718</td>
<td>35.915</td>
</tr>
</tbody>
</table>

* Indicates items selected for the MNA-SF.

(mid-arm circumference), No 5 (institutionalized), No 6 (takes >3 drugs), No 11 (has <3 meals/day), No 13 (has <2 fruits or vegetables/day). The six items with the highest sensitivity and overall accuracy were selected for the MNA-SF. A seventh item with good apparent accuracy, No 17 (self-rated malnourishment), was excluded because of a relatively large proportion (32/155: 21%) of “don’t know” answers. Item No 3 (calf circumference) was not quite as sensitive or accurate as item No 9 and also had the problem of requiring some additional examination time and calculation. Item No 18 (self-rated health) was the next best item, but again was not as good as item No 9 and seemed to have less face validity. These procedures fortunately converged on the same set of six items as possible candidates for inclusion in the MNA-SF.

Figure 1 shows the correlation between the MNA-SF and the full MNA. As can be seen, the correlation is remarkably high ($r = 0.969$). Using a cutpoint for the MNA-SF of ≤ 10/14 as an indicator of possible malnutrition, all of the persons identified on the MNA as “malnourished” (MNA score <17) would have been picked up, as would all but two of those identified as “at risk” (MNA score 17–23). Thus, using the full MNA as the gold standard and setting the threshold at ≤10/14, the sensitivity of the MNA-SF for identifying “malnourished” or “at
Fig. 1. Correlation between the six-item MNA-SF and the full 18-item MNA.

The correlation between the MNA-SF and serum albumin level is quite high ($r = 0.679$) – almost as high as that between albumin and the full MNA ($r = 0.699$). Moreover, the ability of the MNA-SF to identify subnormal albumin ($< 3.5 \text{ g/dl}$) is remarkably good. All but two persons with a subnormal albumin concentration were identified among the persons with MNA-SF scores of $\leq 10/14$. Again, this gives an overall predictive accuracy of 98.7% in this population.

Table 3 shows the relative predictive accuracy of the full MNA and the MNA-SF with regard to recognizing persons with clinically determined definite or possible malnutrition, as well as persons with a serum albumin of $< 3.5 \text{ g/dl}$. With regard to predicting persons with clinical malnutrition, the full MNA has an overall predictive accuracy of 97.1%, versus a virtually indistinguishable 96.4% for the MNA-SF.
As with any screening test, choice of threshold is crucial in determining the diagnostic characteristics. There is usually an inverse relation between sensitivity and specificity, and as the threshold for “normality” is lowered (that is, to include more persons as “abnormal”) the sensitivity of the test increases, while the specificity decreases. This can be seen visually in Figure 2, and it can be seen that for the MNA-SF, the optimal threshold appears to be \(10/14\) as indicating “normal” or low risk.

Based on the above data, it seems very likely that the MNA-SF will be a good first step of a two-step screening strategy for nutritional problems, which would be followed by administration of the full MNA to persons identified as potentially at risk by the MNA-SF. Assuming the population screened was similar to the largely frail one employed here, this two-step strategy would result in a 34% reduction in persons needing to receive the full MNA, since 34% of the Toulouse 1991 population had a MNA-SF score of \(10/14\). If, as is likely, the population to be screened in many geriatric settings was more independent, then a much higher proportion of people would be screened out by the MNA-SF, and the savings would be substantially greater.
Fig. 2. Choosing a threshold (cutpoint) on the MNA-SF for identifying persons with nutritional problems.

Discussion

Assuming that screening for nutritional problems is a worthwhile strategy, it is apparent that the MNA can successfully uncover these problems, being highly correlated with both clinical assessment of nutritional status and with objective indicators such as serum albumin. However, in view of the large amount of information that needs to be collected in the course of the CGA, or even the geriatric screening process, a shorter form than the full MNA, if still accurate, would have real advantages. In this data analysis, we have been able to reduce the size and complexity of the MNA into a shortened form (MNA-SF), which retains most of the beneficial qualities of the original version, at least when the original population is examined. The resultant MNA-SF has six questions (as compared with the original 18) with a score range of 0 to 14, and eliminates most of the most time-consuming items (for example, anthropometric measurements) and questions asking for subjective judgments. It can be administered in approximately 3 min, as opposed to the estimated 10 min for the full MNA. It has a high correlation with both the full MNA ($r = 0.969$) and with serum albumin ($r = 0.679$ vs. 0.699 for the full MNA), and has high concordance with global assessment of nutritional status as being "malnourished vs. normal" (96.4% concordance vs. 97.1% for the
full MNA). Moreover, using a cutpoint of 10/14, the MNA-SF has a sensitivity of 97.9%, a specificity of 100%, and an overall predictive accuracy of 98.7% for predicting malnutrition, using the full MNA as the gold standard.

Some caution is needed before extrapolating these data. First, the MNA-SF was developed and tested on the same population in which the original MNA was developed and tested. The MNA-SF needs to be validated on other populations. While it is likely that its screening characteristics will be even better among lower risk subjects, this speculation must be verified.

A major underlying assumption is that patients found to be malnourished or at risk will be able to benefit from being detected. In reality, it is unclear what proportion of malnourished persons discovered would really be able to benefit by being worked up and/or given nutritional counselling and supplementation, and what proportion are malnourished because of an underlying nonreversible medical condition. To answer this question, we are in desperate need of intervention studies. However, most clinicians in this area are not waiting for the study results, so convincing is the argument for attempting intervention in persons at clear nutritional risk. While treatment will only correct a subgroup of persons discovered to have malnutrition, unless these persons are discovered – probably through a screening or assessment process – this subgroup, regardless of size, is unlikely to benefit.

We must identify a rational strategy for providing nutritional screening in the population. The MNA can certainly be an important part of this strategy. Based on data presented here, it appears that the nutritional screening and assessment can be made even more efficient, and kept just as effective, by using this MNA-SF as the first step in a two-step screening process, which would then be followed up by using the full MNA or other techniques to help confirm the diagnosis and help plan specific treatment for appropriate persons. This strategy must next be tested by first validating the diagnostic characteristics of the MNA-SF itself on different populations, and then, if validated, studying this two-step strategy in clinical trials.

References


**Discussion**

*Dr. Morley:* We have also seen that the MNA is very useful in preoperative evaluation, identifying those people who are malnourished but otherwise look healthy. We have seen that the MNA may play a prospective role in telling us what will happen in the nursing home. We have seen in intervention studies that the MNA may be useful to follow people, and we see that it taps into the cost of care in hospitals. So we have covered a large number of areas. Dr. Rubenstein, I’ll ask you the first question. When you look at the short form of the MNA, how does it compare with the nutritional screening index that has been developed by another nutritional company in the USA? It seems very similar and could almost be used as a pre-screen. Have you had an opportunity to compare your short form of the MNA with the nutritional screening index?

*Dr. Rubenstein:* No, I haven’t compared those two.

*Dr. Morley:* But how do you feel about the nutritional screening index as a tool that people can use in the community, before they get to a dietician or anyone else? It has a high sensitivity, but very poor specificity. It looks as though your short form may have better specificity and could be given to people to self-report.

*Dr. Rubenstein:* I really need to compare the two side by side. I think that the other screening test, which was not based on a more comprehensive assessment, may have omitted two or three items that are important. But I would like to check on that.

*Dr. Schweizer:* You mentioned that the upside of this is that you only need 2 min instead of 10 to carry it out. This may of course decrease or overcome the reluctance of people to use such assessments, as mentioned by Dr. Zulian. The downside, of course, is that you will lose a lot of information that could help you afterwards to design an intervention and to help those patients individually. You say that for two thirds of the patients who would maybe need such an intervention, you would add the full MNA. But my quick calculation shows that the resulting time gain is not very impressive. For 100 patients you would need 1,000 min for a full MNA, and with the combination of short form for all and a full MNA for two thirds, you would need nearly 900 min. Can you comment on that?

*Dr. Rubenstein:* Your point is well taken. For the frail population there is not so much saving if the two-step strategy is applied. The real savings come with its use as a screening test in a general elderly population. In that setting – and we’ll see how it holds up in the New Mexico data and in some of the other databases – the benefits are the major time savings.
Dr. Vellas: An important point in geriatric assessment is that we don’t have to use the same tool in healthy or frail elderly people. For a nutritional assessment, we need to use the MNA for frail elderly people, but for healthy elderly people, a short form test would be satisfactory.

Dr. Arnaud-Battandier: Following up on that, we find that in France general practitioners said they don’t have time to do the MNA in all their elderly patients in their offices. So we are currently testing a list of 12 items which can be answered in 1 min: questions like do they have a disease, have they lost some weight, or do they live alone, etc., and if two of the risk factors are present, the doctor should do the full MNA.

Dr. Camilla: I found Dr. Rubenstein’s intervention very useful. We really need an easy tool for primary care physicians to use in the primary care situation, and we need to be able to educate them into implementing a more comprehensive evaluation.