Randomized Controlled Drug Trials on Very Elderly Subjects: Descriptive and Methodological Analysis of Trials Published Between 1990 and 2002 and Comparison With Trials on Adults

Jean-Laurent Le Quintrec, 1 Caroline Bussy, 2 Jean-Louis Golmard, 3 Christian Hervé, 4 Alain Baulon, 1 and François Piette 5

1Service de Gériatrie I, Hôpital Sainte-Péline/Rossini, Paris, France.
2Service de Gériatrie 3, Hôpital C. Foix, Ivry sur Seine, France.
3Département de Biostatistiques, Faculté Pitié-Salpetrière, Paris, France.
4Laboratoire d’Ethique Médicale, Université Paris 5, France.
5Service de Médecine Interne, Hôpital C. Foix, Ivry sur Seine, France.

Background. Very elderly subjects (VES; aged 80 years or older) constitute a special population as they frequently present multiple diseases (polypathology). Results from trials on general adult populations therefore cannot be extrapolated to VES. We performed a census of randomized controlled trials (RCT) on VES published between 1990 and 2002, and carried out a descriptive and methodological analysis of these RCT/VES, comparing them with matched RCT on general adult populations (control RCT, RCT/C).

Methods. We searched for RCT/VES in two international databases (EMBASE and MEDLINE) and then manually. RCT/C were matched to RCT/VES for disease area and year of publication. The methodological quality of each RCT was assessed with Chalmers’ scale.

Results. We identified 84 RCT/VES, 63 of which were conclusive and 21, inconclusive. Subjects were institutionalized in 48 RCT, and community dwelling in 11 RCT (unspecified in 25 RCT). Efficacy was the main criterion in 75 RCT; tolerance in 9 RCT. Twenty-six RCT were published by geriatrics journals, and 58 by general medical journals. The RCT/VES covered most of the disease areas of geriatrics. The 84 RCT/VES had a mean methodological quality score of 0.578 ± 0.157. The matched 84 RCT/C had a mean methodological quality score of 0.592 ± 0.116 (p = .466). The methodological quality score of RCT/VES increased with the number of included subjects (p = .004) and the year of publication (p = .001).

Conclusions. The methodological quality of RCT/VES is equivalent to that of RCT in general adult populations. Nevertheless, RCT/VES remain very scarce, and neglect certain diseases. RCT/VES and the inclusion of very elderly subjects in RCT on adults should be strongly encouraged.

THE number of elderly people is increasing in developed countries. Within this population, the proportion of very elderly subjects (VES; 80 years or older) is increasing the most rapidly. A large proportion of VES differ from other adults and elderly people in suffering from multiple diseases (polypathology) which modify drug metabolism, efficacy, and side effects (1,2). These subjects therefore need specific medical care; specific medical research focusing on this population is necessary. However, little medical research is carried out on VES, who are commonly excluded from randomized controlled trials (RCT) solely on the basis of their age (3,4). This exclusion of VES has been criticized in research on cancer (5–9), on the treatment of myocardial infarction (1,10), and on Parkinson’s disease (2). RCT specifically devoted to VES may overcome this problem, but such trials are rare: Bugeja (3) identified only 18 RCT on subjects aged 75 years or over among 490 published RCT. Public health authorities have been trying hard to promote clinical research on VES since the 1980s in the United States and the 1990s in Europe. However, specific ethical (4,11) and methodological (12) concerns limit the development of RCT on VES. This study had two specific goals: 1) to produce an inventory of published RCT on VES and 2) to analyze these RCT, assess their methodological quality, and compare them with RCT on general adult populations.

METHODS
We performed a census of RCT on drug therapy in subjects with a mean age over 80 years, published between January 1990 and December 2002. Our search was not limited to the English language. We began by searching two electronic databases: EMBASE and MEDLINE. We selected a large group of key words (randomized controlled trials, randomized controlled trials, double blind method, prospective studies, comparative studies, crossover studies, controlled clinical trials), which were used as “medical subject headings” and/or “text words.” We then manually checked the references of each selected article.
We assessed the following characteristics of the RCT on VES (RCT/VES): aim of the trial (efficacy or tolerance), type of endpoint (major clinical or surrogate), result (positive or inconclusive), age, and autonomy of included subjects, disease areas, type of journal publishing the RCT (geriatrics journal or general medical journal), number of included subjects, and methodological quality (as assessed with Chalmers' scale). A major clinical endpoint was defined as the “true endpoint” (13) for which the tested drug was developed, e.g., decrease in frequency of infections for prophylactic anti-infection drugs, decrease in frequency of fractures for antiosteoporotic drugs. Conversely, a surrogate endpoint was defined as a nonclinical and/or nonmajor endpoint, e.g., increase in biological immune markers for prophylaxis of infections, or increase in bone mineral density for antiosteoporotic drugs.

We then searched for control RCT (RCT/C) on subjects with a mean age < 80 years. Each RCT/C was randomly selected from all RCT on adults (18 years or over) published in the same disease area and during the same year as the RCT/VES to which they were matched. For each RCT/C, we assessed aim of the trial, type of endpoint, results, mean age, number of included subjects, and methodological quality (using Chalmers’ scale).

We assessed the methodological quality of RCT using Chalmers’ scale (14). This scale, which was not specifically designed for geriatric RCT, is very extensive (15). It comprises 27 items grouped into three sections: the first section (14 items) assesses the design of the trial and is scored out of 60; the second section (9 items) assesses statistical analysis and is scored out of 30; and the third section (4 items) assesses presentation of the results and is scored out of 10. The final score obtained is therefore between 0/100 (lowest quality) and 100/100 (highest quality) and has been scaled to give a number between 0 and 1. Each article was assessed by two blinded independent geriatricians. The final score was obtained by consensus, as suggested by Chalmers (14).

Statistical analyses were performed in two steps. In the first, each group of trials—RCT/VES and RCT/C—was studied separately, to assess the relationships between methodological quality score and aim of the trial, type of endpoint, result of the trial, type of journal, number of subjects, and year of publication. All the comparisons were performed using nonparametric tests (Wilcoxon tests) for qualitative variables and Spearman correlation coefficient tests for quantitative variables. In the second step, RCT/VES and RCT/C were compared by means of pair-wise Student’s t tests for quantitative variables and McNemar tests for qualitative variables. The SAS 8.1 package (SAS Institute, Cary, NC) was used for all statistical calculations. We used a significance threshold of p = .05 for all tests.

**RESULTS**

We identified 84 RCT/VES on drug therapy published between January 1990 and December 2002. These RCT/VES covered most of the disease areas of geriatrics (Table 1). The aim of the RCT concerned drug efficacy in 75 RCT, and drug tolerance (Table 2) in nine. The endpoint of the trial was major clinical in 65 RCT, and surrogate in 19 RCT. Among the 84 RCT/VES, 63 led to a conclusive result, and 21 were inconclusive.

The mean age of included subjects was between 80.0 and 88.0 years (overall mean age of all subjects included: 83.0 ± 2.0 years). Subjects were community dwelling in 11 RCT, and were institutionalized in 48 RCT (not specified in 25 RCT).

Fifty-eight of the RCT were published by general medical journals, and 26 by geriatrics journals. The number of included subjects was between 7 and 5125 (mean ± SD: 264 ± 715) and did not differ significantly according to the aim of the trial (p = .259), type of endpoint (p = .556), result (p = .430), or type of journal (p = .630).

The mean methodological quality score of the 84 RCT/VES was 0.578 ± 0.157 and did not differ significantly according to the aim of the trial (p = .558), type of endpoint (p = .101), result (p = .098), or type of journal (p = .654). The methodological quality score of RCT/VES increased with the number of included subjects (p = .004) and with the year of publication (p = .001).

In the 84 RCT/C, the aim concerned drug efficacy in 73 RCT and drug tolerance in 11. The endpoint was major clinical in 62 RCT and surrogate in 22. Among the 84 RCT/C, 62 led to a conclusive result, and 22 were inconclusive. The mean age of included subjects was between 22.3 and 79.6 years (overall mean age of all subjects included: 57.2 ± 1.8 years). There was an upper age limit in 18 RCT/C.
The number of included subjects in the 84 RCT/C was between 10 and 21,106 (mean ± SD = 530 ± 2482). The number of included subjects was significantly higher in RCT/C with a major clinical endpoint than in those with a surrogate endpoint \((p < .0001)\) but did not differ significantly according to the aim of the trial \((p = .600)\), result \((p = .202)\), or type of journal \((p = .170)\).

The mean methodological quality score of the 84 RCT/C was 0.592 ± 0.116. The methodological quality score was significantly higher in RCT/C with a major clinical endpoint than in those with a surrogate endpoint \((p < .0001)\), but did not differ significantly according to the aim of the trial \((p = .883)\), result \((p = .497)\), or type of journal \((p = .210)\). The methodological quality score of the RCT/C increased with the number of included subjects \((p = .0007)\) and with the year of publication \((p = .006)\).

Comparisons between RCT/VES and RCT/C revealed no statistically significant difference for study protocol (Table 3). The number of included subjects was similar in RCT/VES and RCT/C \((p = .757)\). Finally, the methodological quality was equivalent in RCT/VES and RCT/C \((p = .466)\).

**DISCUSSION**

This census of RCT/VES published in medical journals since 1990 identified only 84 RCT/VES. This number is very small, given that some 50,000 RCT on humans were published during this period, according to the databases searched. This small number of RCT/VES does not seem to be an underestimate, in that these 84 RCT/VES are restricted to drug therapy and cover most of the main subject headings of geriatrics (Table 1). Furthermore, because electronic databases are thought to introduce bias favoring articles in the English language (16), we supplemented our database search with a manual search. Our final search strategy could be estimated to be 75% exhaustive (17).

We searched for RCT/C using the same search strategy as for RCT/VES. RCT/C were matched with RCT/VES for year of publication, to avoid any discrepancy due to possible improvements in methodological quality between 1990 and 2002. RCT/C were also matched with RCT/VES for disease area, to allow valuable comparisons between RCT/VES and RCT/C. This strategy also eliminated bias due to journals, in that we searched for RCT/C not only in the journals that published RCT/VES, but in all journals dealing with the same disease area. The overall mean age of all subjects included in the 84 RCT/C was very different from that of all subjects in the 84 RCT/VES (57.2 ± 1.8 years vs 83.0 ± 2.0 years), and there was therefore a minimal age distribution overlap between the two groups of RCT.

Most of the RCT/VES were designed to determine drug efficacy, which is the primary goal for RCT. Tolerance is also a major concern in geriatrics, because the risk/benefit ratio of treatments in VES differs greatly from that in adults as a whole (10). This difference precludes the extrapolation of results obtained in populations of adults. In our study, only 9 RCT/VES and 11 RCT/C aimed to assess drug tolerance. The proportion of RCT/VES aimed at assessing drug tolerance would therefore seem to be insufficient, given that the prevalence of multiple chronic conditions and subsequent polymedication, generating numerous side effects, increases with age (18). There is also a need for RCT identifying the minimum effective dose of some drugs used in VES, especially drugs for Parkinson’s disease, diabetes, hypertension, and coronary disease (19).

The endpoint of RCT should be clinical, as much as possible, because only major clinical endpoints can demonstrate the real clinical efficacy of the drug tested. In some disease areas (e.g., osteoporosis) and for preventive treatments (e.g., cardiovascular diseases and cancer), the clinical endpoint requires long-term follow-up and a large number of included subjects; neither of these conditions is easy to fulfill with VES. Surrogate endpoints may be used in such cases, provided that the surrogate endpoint has been shown to be linked to the corresponding clinical endpoint. Nineteen of the 84 RCT/VES were based on surrogate endpoints. Most of these 19 RCT/VES dealt with disease prevention—for osteoporosis (7 RCT), infections (7 RCT), nutritional problems (3 RCT), and serum lipid concentrations (1 RCT). The proportion of trials with surrogate endpoints was similar in RCT/VES and RCT/C. Given the difficulties involved in carrying out RCT/VES with a major clinical endpoint, geriatricians cannot be considered to “overuse” surrogate endpoints.

Twenty-one RCT/VES were inconclusive: this is an important finding because it was thought until recently that pharmaceutical companies and editorial boards avoided publishing inconclusive trials (20). The distribution of inconclusive RCT/VES between disease areas was homogeneous and did not differ from that of conclusive RCT/VES. The proportion of inconclusive trials among RCT/VES was similar to that of inconclusive trials among RCT/C. The publication of inconclusive trials therefore seems to be becoming more common for studies of both geriatric and general adult populations; this trend should be encouraged.

Despite the high mean age of included subjects, 11 RCT/VES were devoted to community-dwelling subjects. Most of these 11 RCT dealt with disease prevention (5 RCT) or preservation of autonomy (4 RCT), both of which can only effectively be studied in autonomous subjects living in the community. In 25 RCT/VES, the level of autonomy of the included subjects was not specified. This lack of detail is a particularly serious omission in RCT in geriatrics, because VES constitute a very heterogeneous population. In this population, age alone is insufficient for the assessment of autonomy and health because of differential aging: some people aged 80 or older live in the community like other adults, whereas others suffer multiple diseases and live in nursing homes. It is therefore essential for both scientific and ethical reasons to report the place of residence or

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>RCT/VES ((N = 84))</th>
<th>RCT/C ((N = 84))</th>
<th>(p) Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aim: efficacy</td>
<td>75</td>
<td>73</td>
<td>.999</td>
</tr>
<tr>
<td>Endpoint: major clinical</td>
<td>65</td>
<td>62</td>
<td>.117</td>
</tr>
<tr>
<td>Result: conclusive</td>
<td>63</td>
<td>62</td>
<td>.141</td>
</tr>
</tbody>
</table>

*Chi-square test.
lifestyle in RCT/VES, because these factors define the autonomy or polypathology of the VES included in the trial.

Two major disease areas are neglected by RCT/VES: Parkinson’s disease and cancer. For Parkinson’s disease, Mitchell and colleagues (2) conducted a systematic search of all RCT published between 1966 and 1996. Of the 112 articles they identified, 42 included subjects older than 75 years, and only eight adequately reported the age groups of the included subjects. In our search, we found no RCT on drugs to treat Parkinson’s disease in subjects aged 80 or older published between 1990 and 2002. The exclusion of elderly subjects from cancer research has been criticized since the 1980s. In 1987, Begg and Carbone (5) showed that patients aged 70 years or older had rates of response and toxicity similar to those in younger adults, and agreed with the National Cancer Institute statement that age is not a criterion for the exclusion of patients. In 1994, Trimble and colleagues (6) found that “with respect to cancer incidence, older patients (65 years and older) are underrepresented in cancer treatment trials.” Montfardini and colleagues (7), in phase II trials including 2344 patients, found that 22% of patients were aged 65 years or older and only 8% of patients were aged 70 years or older.

Fifty-eight of the 84 RCT/VES were published by general medical journals. Clinical trials on geriatric topics are no longer confined to geriatrics journals but are now considered interesting by the editorial boards of general medical journals. The RCT/VES published by geriatrics journals were mainly devoted to “specific” geriatric disease areas such as urinary incontinence, Alzheimer’s disease, nutritional problems, and immune function.

The mean methodological quality score of the 84 RCT/VES was 0.578 ± 0.157. This score is very different from the perfect score of 1 for the Chalmers’ scale. This low mean score may be due to the nature of the Chalmers’ scale, which was not specifically designed for geriatric research. However, there are two main reasons why this is unlikely to be the case. First, although clinical research on VES certainly requires some adaptation of traditional methodologies and modification of expectations (12), RCT/VES are above all RCT, and therefore need to satisfy the same methodological quality criteria as RCT on any other subject. Second, the mean score of the RCT/C (0.592 ± 0.116) was very similar to that of RCT/VES (0.578 ± 0.157) (p = .466): the same search applied to RCT on VES and matched RCT on general adult populations gave similar results. The Chalmers’ scale is therefore not biased in geriatric clinical trials.

This analysis of RCT/VES published between 1990 and 2002 highlights the rarity of RCT/VES. Clinical research on VES involves specific difficulties, which must be resolved if RCT are to be conducted in such populations. The principal difficulty is the great heterogeneity of VES, due to differential aging and multiple diseases. The recruitment strategy should therefore be broad, to avoid selecting subjects presenting only one disease; these subjects may not constitute a representative sample of the target population for the drug studied (18). So, inappropriate inclusion criteria may result in nonrepresentative samples of patients, as has been demonstrated in cancer research (9,21–23). There should therefore be very few exclusion criteria, and these should be based on ethical or pathophysiological grounds (1,9,10,23, 24,25): inability to obtain informed consent from subject; imminent death of subject; contraindication of subject to the drug studied. Age itself should no longer be considered an exclusion criterion.

The second difficulty is that patients or their proxies are more reluctant to participate in RCT if the patient is very old and suffers from multiple diseases. This reluctance is often shared by the patient’s personal physician (26). These difficulties may result in the recruitment of less disabled patients, who do not constitute a representative sample of the population. Such bias has been demonstrated in studies of elderly patients with acute myocardial infarction or lymphoma (21).

The third difficulty is the limited life expectancy of VES. The number of deaths should be estimated before enrollment is begun. This restricted life expectancy makes it necessary to choose an appropriate endpoint: morbidity and mortality endpoints are often not suitable for VES (2), especially in clinical trials dealing with preventive or symptomatic treatments. In these cases, indices of quality of life and activities of daily living are required (27).

Frequent cognitive and sensory deterioration in VES is also a difficulty, resulting in specific problems, such as poor compliance during the RCT (28), and may necessitate changes to the number and circumstances of clinical follow-up appointments (29). These difficulties are particularly great in clinical research focusing on Alzheimer’s disease patients; this research should be restricted to drugs aimed at Alzheimer’s disease treatment.

Our descriptive and methodological analysis showed a) that RCT/VES published since 1990 have a methodological level equivalent to that of RCT on general adult populations, and b) that methodological quality has increased over time. This result is very encouraging for geriatricians. Clinical research in geriatrics should now be aimed at performing RCT on VES, who constitute a growing population of patients suffering from cancer (30), Parkinson’s disease (2), and cardiovascular disease (31), in particular.

Furthermore, researchers testing a new treatment on adults must include VES in RCT. As stated by the International Conference of Harmonization (ICH) Geriatric Guidelines (32), VES must be included in sufficient numbers in all RCT testing drugs “that are likely to have significant use in the elderly . . . The older the population likely to use the drug, the more important it is to include the very old.” Geriatricians, researchers, regulatory authorities, and pharmaceutical companies now have to translate these guidelines into facts.

ACKNOWLEDGMENTS

Address correspondence to Jean-Laurent Le Quintrec, Hôpital Sainte-Périche/Rossini, Service de Gerontologie I, 11, rue Chardon Lagache, F 75016 Paris, France. E-mail: muriel.palomares@spR.hop-paris.fr

REFERENCES


