The CONSORT Statement: Revised Recommendations for Improving the Quality of Reports of Parallel-Group Randomized Trials

David Moher, MSc
Kenneth F. Schulz, PhD, MBA
Douglas Altman, DSc
for the CONSORT Group

To comprehend the results of a randomized controlled trial (RCT), readers must understand its design, conduct, analysis, and interpretation. That goal can be achieved only through complete transparency from authors. Despite several decades of educational efforts, the reporting of RCTs needs improvement. Investigators and editors developed the original CONSORT (Consolidated Standards of Reporting Trials) statement to help authors improve reporting by using a checklist and flow diagram. The revised CONSORT statement presented in this article incorporates new evidence and addresses some criticisms of the original statement.

The checklist items pertain to the content of the Title, Abstract, Introduction, Methods, Results, and Comment. The revised checklist includes 22 items selected because empirical evidence indicates that not reporting the information is associated with biased estimates of treatment effect or because the information is essential to judge the reliability or relevance of the findings. We intended the flow diagram to depict the passage of participants through an RCT. The revised flow diagram depicts information from 4 stages of a trial (enrollment, intervention allocation, follow-up, and analysis). The diagram explicitly includes the number of participants, according to each intervention group, included in the primary data analysis. Inclusion of these numbers allows the reader to judge whether the authors have performed an intention-to-treat analysis.

In sum, the CONSORT statement is intended to improve the reporting of an RCT, enabling readers to understand a trial's conduct and to assess the validity of its results.

JAMA. 2001;285:1987-1991 www.jama.com

A REPORT OF A RANDOMIZED CONtrolled trial (RCT) should convey to the reader, in a transparent manner, why the study was undertaken and how it was conducted and analyzed. For example, a lack of adequately reported randomization has been associated with bias in estimating the effectiveness of interventions.[1,2] To assess the strengths and limitations of an RCT, readers need and deserve to know the quality of its methods. Despite several decades of educational efforts, RCTs still are not being reported adequately.[3-6] For example, a review of 122 recently published RCTs that evaluated the effectiveness of selective serotonin reuptake inhibitors as first-line management strategy for depression found that only 1 (0.8%) article described randomization adequately.[5] Inadequate reporting makes the interpretation of RCT results difficult if not impossible. Moreover, inadequate reporting borders on unethical practice when biased results receive false credibility.

HISTORY OF CONSORT

In the mid 1990s, 2 independent initiatives to improve the quality of reports of RCTs led to the publication of the CONSORT (Consolidated Standards of Reporting Trials) statement,[7] which was developed by an international group of clinical trialists, statisticians, epidemiologists, and biomedical editors. CONSORT has been supported by a growing number of medical and health care journals[8-11] and editorial groups, including the International Committee of Medical Journal Editors[12] (ICMJE, also known as the Vancouver Group), the Council of Science Editors (CSE), and

See also pp 1992, 1996, 2000, and 2006.

Author Affiliations: University of Ottawa, Thomas C. Chalmers Centre for Systematic Reviews, Ottawa, Ontario (Mr Moher); Family Health International and Department of Obstetrics and Gynecology, School of Medicine, University of North Carolina at Chapel Hill (Dr Schulz); and ICRF Medical Statistics Group and Centre for Statistics in Medicine, Institute of Health Sciences, Oxford, England (Dr Altman).
Members of the CONSORT Group are listed at the end of this article.
Simultaneous Publication: The revised CONSORT statement also appears in *Annals of Internal Medicine* (2001;134:657-662) and *The Lancet* (2001;357:1191-1194).
Correspondence and Reprints: Leah Lepage, PhD, Thomas C. Chalmers Centre for Systematic Reviews, Children's Hospital of Eastern Ontario Research Institute, Room R235, 401 Smyth Rd, Ottawa, Ontario, Canada K1H 8L1 (e-mail: llepage@uottawa.ca).

the World Association of Medical Editors (WAME). CONSORT is also published in Dutch, English, French, German, Japanese, and Spanish. It can be accessed on the Internet, along with other information about the CONSORT group.[13]

The CONSORT statement comprises a checklist and flow diagram for reporting an RCT. For convenience, the checklist and diagram together are called simply *CONSORT*. They are primarily intended for use in writing, reviewing, or evaluating reports of simple 2-group parallel RCTs.

Preliminary data indicate that the use of CONSORT does indeed help to improve the quality of reports of RCTs.[14,15] In an evaluation of 71 published RCTs in 3 journals in 1994, allocation concealment was reported unclearly in 43 (61%) of the trials.[14] Four years later, after these 3 journals required that authors reporting an RCT use CONSORT, the proportion of articles in which allocation concealment was reported unclearly had decreased to 30 of 77 (39%; mean difference, −22%; [95% confidence interval, −38% to −6%]).[14]

The usefulness of CONSORT is enhanced by continuous monitoring of the biomedical literature; this monitoring allows CONSORT to be modified depending on the merits of maintaining or dropping current items and including new items. For example, when Meinert[16] observed that the flow diagram did not provide important information about the number of participants who entered each phase of an RCT (enrollment, treatment allocation, follow-up, and data analysis), the diagram was able to be modified to accommodate the information. The checklist is similarly flexible.

This iterative process makes the CONSORT statement a continually evolving instrument. While participants in the CONSORT group and their degree of involvement vary over time, members meet regularly to review the need to refine CONSORT. At the 1999 meeting, participants decided to revise the original statement. This report reflects changes determined by consensus of the CONSORT group, partly in response to emerging evidence on the importance of various elements of RCTs.

REVISION OF THE CONSORT STATEMENT

Thirteen members of the CONSORT group met in May 1999 with the primary objective of revising the original CONSORT checklist and flow diagram, as needed. The group discussed the merits of including each item in the light of current evidence. As in developing the original CONSORT statement, our intention was to keep only those items deemed fundamental to reporting standards for an RCT. Some items not considered essential may well be highly desirable and still should be included in an RCT report even though they are not included in CONSORT. Such items include approval of an institutional ethical review board, sources of funding for the trial, and a trial registry number (eg, the International Standard Randomized Controlled Trial Number [ISRCTN]) used to register the RCT at its inception.[17]

Shortly after the meeting, a revised version of the checklist was circulated to the group for additional comments and feedback. Revisions to the flow diagram were similarly made. All these changes were discussed when CONSORT participants met in May 2000, and the revised statement was finalized shortly afterward.

The revised CONSORT statement includes a 22-item checklist (TABLE) and a flow diagram (FIGURE). Its primary aim is to help authors improve the quality of reports of simple 2-group parallel RCTs. However, the basic philosophy underlying the development of the statement can be applied to any design. In this regard, additional statements for other designs will be forthcoming from the group.[13] CONSORT can also be used by peer reviewers and editors to identify reports with inadequate description of trials and those with potentially biased results.[1,2]

During the 1999 meeting, the group also discussed the benefits of developing an explanatory document to enhance the use and dissemination of CONSORT. The document is patterned on reporting of statistical aspects of clinical research[18] and was developed to help facilitate the recommendations of the ICMJE's *Uniform Requirements for Manuscripts Submitted to Biomedical Journals*. Three members of the CONSORT group, with assistance from members on some checklist items, drafted an explanation and elaboration document. That document[19] was circulated to the group for additions and revisions and was last revised after review at the latest CONSORT group meeting.

CHANGES TO CONSORT

(1) In the revised checklist, a new column for "paper section and topic" integrates information from the "subheading" column that was contained in the original statement.

(2) The "Was it reported?" column has been integrated into a "reported on page #" column, as requested by some journals.

(3) Each item of the checklist is now numbered and the syntax and order have been revised to improve the flow of information.

(4) "Title" and "Abstract" are now combined in the first item.

(5) While the content of the revised checklist is similar to the original, some items that previously were combined are now separate. For example, authors had been asked to describe "primary and secondary outcome(s) measure(s) and the minimum important difference(s), and indicate how the target sample size was projected." In the new version, issues pertaining to outcomes (item 6) and sample size (item 7) are separate, enabling authors to be more explicit about each. Moreover, some items request additional information. For example, for outcomes (item 6) authors are asked to report any methods used to enhance the quality of measurements, such as multiple observations.

(6) The item asking for the unit of randomization (eg, cluster) has been dropped because specific checklists have

been developed for reporting cluster RCTs[20] and other design types[13] since publication of the original checklist.

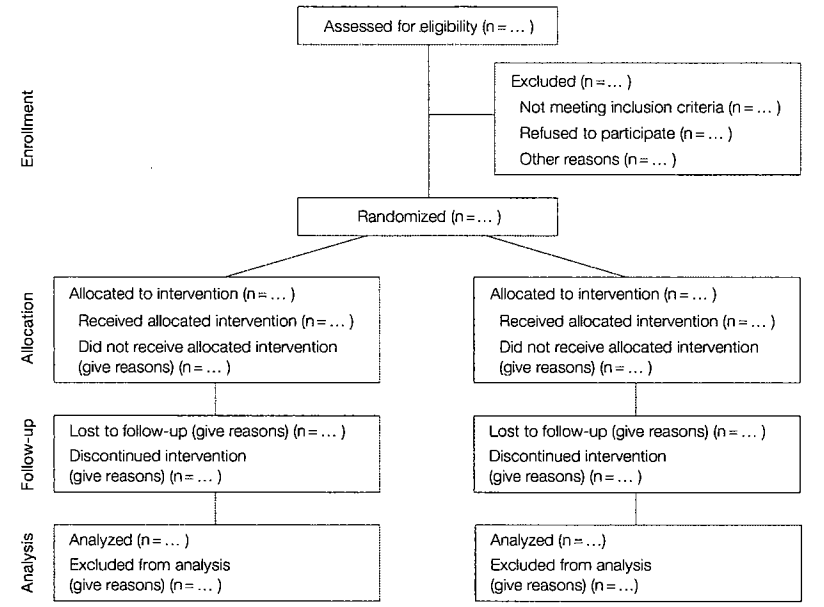
(7) Whenever possible, new evidence is incorporated into the revised checklist. For example, authors are asked to be explicit about whether the analysis reported is by intention-to-treat (item 16). This request is based in part on the observations[21] that authors do not adequately describe and apply intention-to-treat analysis and reports that not providing this information are less likely to provide other relevant information, such as loss to follow-up.[22]

(8) The revised flow diagram depicts information from 4 stages of a trial (enrollment, intervention allocation, follow-up, and analysis). The revised diagram explicitly includes the number of participants, according to each intervention group, included in the primary data analysis. Inclusion of these numbers lets the reader know whether the authors have performed an intention-to-treat analysis,.[21-23] Because some of the information may not always be

Table. Checklist of Items to Include When Reporting a Randomized Trial

Section and Topic	Item #	Descriptor	Reported on Page #
Title and Abstract	1	How participants were allocated to interventions (eg, "random allocation," "randomized," or "randomly assigned").	
Introduction			
Background	2	Scientific background and explanation of rationale.	
Methods			
Participants	3	Eligibility criteria for participants and the settings and locations where the data were collected.	
Interventions	4	Precise details of the interventions intended for each group and how and when they were actually administered.	
Objectives	5	Specific objectives and hypotheses.	
Outcomes	6	Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (eg, multiple observations, training of assessors).	
Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules.	
Randomization			
Sequence generation	8	Method used to generate the random allocation sequence, including details of any restriction (eg, blocking, stratification).	
Allocation concealment	9	Method used to implement the random allocation sequence (eg, numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.	
Implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups.	
Blinding (masking)	11	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. If done, how the success of blinding was evaluated.	
Statistical methods	12	Statistical methods used to compare groups for primary outcome(s); methods for additional analyses, such as subgroup analyses and adjusted analyses.	
Results			
Participant flow	13	Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome. Describe protocol deviations from study as planned, together with reasons.	
Recruitment	14	Dates defining the periods of recruitment and follow-up.	
Baseline data	15	Baseline demographic and clinical characteristics of each group.	
Numbers analyzed	16	Number of participants (denominator) in each group included in each analysis and whether the analysis was by "intention-to-treat." State the results in absolute numbers when feasible (eg, 10/20, not 50%).	
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each group, and the estimated effect size and its precision (eg, 95% confidence interval).	
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those prespecified and those exploratory.	
Adverse events	19	All important adverse events or side effects in each intervention group.	
Comment			
Interpretation	20	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes.	
Generalizability	21	Generalizability (external validity) of the trial findings.	
Overall evidence	22	General interpretation of the results in the context of current evidence.	

Figure. Flow Diagram of Subject Progress Through the Phases of a Randomized Trial



known and to accommodate other information, the structure of the flow diagram may need to be modified for a particular trial. Inclusion of the participant flow diagram in the report is strongly recommended but may be unnecessary for simple trials, such as those without any participant withdrawals or dropouts.

COMMENT

Specifically developed to guide authors about how to improve the quality of reporting of simple 2-group parallel RCTs, CONSORT encourages transparency in reporting the methods and results so that reports of RCTs can be interpreted both readily and accurately. However, CONSORT does not address other facets of reporting that also require attention, such as scientific content and readability of RCT reports. Some authors, in their enthusiasm to use CONSORT, have modified the checklist.[24] We recommend against such modifications because they may be based on a different process than the one used by the CONSORT group.

The use of CONSORT seems to reduce (if not eliminate) inadequate reporting of RCTs.[14,15] Potentially, the use of CONSORT should positively influence the manner in which RCTs are conducted. Granting agencies have noted this potential relationship and, in at least 1 case,[25] have encouraged grantees to consider in their application how they have dealt with the CONSORT items.

The evidence-based approach used to develop CONSORT also has been used to develop standards for reporting meta-analyses of randomized trials,[26] meta-analyses of observational studies,[27] and diagnostic studies (Jeroen Lijmer, MD, written communication, October 2000). Health economists also have started to develop reporting standards[28] to help improve the quality of their reports.[29] The intent of all these initiatives is to improve the quality of reporting of biomedical research[30] and by doing so to bring about more effective health care.

The revised CONSORT statement will replace the original one in the journals and groups that already support it. Journals that do not yet support CONSORT may do so by registering on the CONSORT Web site.[13] To convey to authors the importance of improved quality in the reporting of RCTs, we encourage supporting journals to reference the revised CONSORT statement and the CONSORT Internet address[13] in their "Instructions to Authors." Because the journals publishing the revised CONSORT statement have waived copyright protection, CONSORT is now widely accessible to the biomedical community. The CONSORT checklist and flow diagram can also be accessed at the CONSORT Web site.[13]

A lack of clarification of the meaning and rationale for each checklist item in the original CONSORT statement has been remedied with the development of the CONSORT explanation and elaboration document,[19] which also can be found on the CONSORT Web site.[13] This document reports the evidence on which the checklist items are based, including the references, which had annotated the checklist items in the previous version. We encourage journals to also include reference to this document in their Instructions to Authors.

Emphasizing the evolving nature of CONSORT, the CONSORT group invites readers to comment on the updated checklist and flow diagram through the CONSORT Web site.[13] Comments and suggestions will be collated and considered at the next meeting of the group in 2001.

Author Contributions: Mr Moher and Drs Schulz and Altman participated in regular conference calls, identified participants, contributed to the CONSORT meetings, and drafted the manuscript. Mr Moher planned the CONSORT meetings, secured funding, invited the participants, and planned the meeting agenda. Members of the CONSORT group attended the meetings and provided input for the revised checklist, flow diagram, and/or text of this article. Mr Moher is the guarantor of this article.

Contributors to this version of the CONSORT statement: Frank Davidoff, MD, *Annals of Internal Medicine* (Philadelphia, Pa); Susan Eastwood, ELS(D), University of California at San Francisco; Matthias Egger, MD, Department of Social Medicine, University of Bristol (Bristol, England); Diana Elbourne, PhD, London School of Hygiene and Tropical Medicine (London, England); Peter Gøtzsche, MD, Nordic Cochrane Centre (Copenhagen, Denmark); Sylvan B. Green, PhD, MD, School of Medicine, Case Western Reserve University (Cleveland, Ohio); Leni Grossman, BA, Merck & Co Inc (Whitehouse Station, NJ); Barbara S. Hawkins, PhD, Wilmer Ophthalmological Institute, Johns Hopkins University (Baltimore, Md); Richard Horton, MB, *The Lancet* (London, England); Wayne B. Jonas, MD, Uniformed Services University of the Health Sciences (Bethesda, Md); Terry Klassen, MD, Department of Pediatrics, University of Alberta (Edmonton); Leah Lepage, PhD, Thomas C. Chalmers Centre for Systematic

Reviews (Ottawa, Ontario); Thomas Lang, MA, Tom Lang Communications (Lakewood, Ohio); Jeroen Lijmer, MD, Department of Clinical Epidemiology, University of Amsterdam (Amsterdam, the Netherlands); Rick Malone, BS, TAP Pharmaceuticals (Lake Forest, Ill); Curtis L. Meinert, PhD, Johns Hopkins University (Baltimore, Md); Mary Mosley, BS, Life Science Publishing (Tokyo, Japan); Stuart Pocock, PhD, London School of Hygiene and Tropical Medicine (London, England); Drummond Rennie, MD, *JAMA* (Chicago, Ill); David S. Riley, MD, University of New Mexico Medical School (Santa Fe); Roberta W. Scherer, PhD, Epidemiology and Preventive Medicine, University of Maryland School of Medicine (Baltimore); Ida Sim, MD, PhD, University of California at San Francisco; and Donna Stroup, PhD, MSc, Epidemiology Program Office, Centers for Disease Control and Prevention (Atlanta, Ga).

Funding/Support: Financial support to convene meetings of the CONSORT group was provided in part by Abbott Laboratories, American College of Physicians, GlaxoWellcome, *The Lancet*, Merck, Canadian Institutes for Health Research, National Library of Medicine, and TAP Pharmaceuticals.

Acknowledgment: The effort to improve the reporting of randomized trials, from its beginnings with the Standards of Reporting Trials (SORT) group to the current activities of the CONSORT group, have involved a large number of people around the globe. We wish to thank Leah Lepage, PhD, for keeping everybody all lined up and moving in the same direction.

REFERENCES

1. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*. 1995;273:408-412.
2. Moher D, Pham B, Jones A, et al. Does the quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*. 1998;352:609-613.
3. Jadad AR, Boyle M, Cunningham C, Kim M, Schachar R. *Treatment of Attention-Deficit/Hyperactivity Disorder*. Hamilton, Ontario: McMaster University; 2000. Evidence Report/Technology Assessment No. 11.
4. Thornley B, Adams CE. Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ*. 1998;317:1181-1184.
5. Hotopf M, Lewis G, Normand C. Putting trials on trial: the costs and consequences of small trials in depression: a systematic review of methodology. *J Epidemiol Community Health*. 1997;51:354-358.
6. Dickinson K, Bunn F, Wentz R, Edwards P, Roberts I. Size and quality of randomised controlled trials in head injury: review of published studies. *BMJ*. 2000; 320:1308-1311.
7. Begg CB, Cho MK, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA*. 1996;276:637-639.
8. Freemantle N, Mason JM, Haines A, Eccles MP. CONSORT: an important step toward evidence-based health care. *Ann Intern Med*. 1997;126:81-83.
9. Altman DG. Better reporting of randomized controlled trials: the CONSORT statement. *BMJ*. 1996; 313:570-571.
10. Schulz KF. The quest for unbiased research: randomized clinical trials and the CONSORT reporting guidelines. *Ann Neurol*. 1997;41:569-573.
11. Huston P, Hoey J. CMAJ endorses the CONSORT statement. *CMAJ*. 1996;155:1277-1279.
12. Davidoff F. News from the International Committee of Medical Journal Editors. *Ann Intern Med*. 2000;133:229-231.
13. CONSORT Web site. Available at: http://www.consort-statement.org. Accessibility verified March 14, 2001.
14. Moher D, Jones A, Lepage L, for the CONSORT Group. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA*. 2001;285:1992-1995.
15. Egger M, Jüni P, Bartlett C, for the CONSORT Group. Value of flow diagrams in reports of randomized controlled trials. *JAMA*. 2001;285:1996-1999.
16. Meinert CL. Beyond CONSORT: need for improved reporting standards for clinical trials. *JAMA*. 1998;279:1487-1489.
17. Chalmers I. Current controlled trials: an opportunity to help improve the quality of clinical research. *Curr Control Trials Cardiovasc Med*. 2000;1:3-8.
18. Bailer JC III, Mosteller F. Guidelines for statistical reporting in articles for medical journals: amplifications and explanations. *Ann Intern Med*. 1988;108: 266-273.
19. Altman DG, Schulz KF, Moher D, et al, for the CONSORT group. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001;134:663-694.
20. Elbourne DR, Campbell MK. Extending the CONSORT statement to cluster randomised trials: for discussion. *Stat Med*. 2001;20:489-496.
21. Hollis S, Campbell F. What is meant by intention-to-treat analysis? survey of published randomized controlled trials. *BMJ*. 1999;319:670-674.
22. Ruiz-Canela M, Martinez-Gonzalez MA, de Irala-Estevez J. Intention-to-treat analysis is related to methodological quality. *BMJ*. 2000;320:1007.
23. Lee YJ, Ellenberg JH, Hirtz DG, Nelson KB. Analysis of clinical trials by treatment actually received: is it really an option? *Stat Med*. 1991;10:1595-1605.
24. Bentzen SM. Towards evidence based radiation oncology: improving the design, analysis, and reporting of clinical outcome studies in radiotherapy. *Radiother Oncol*. 1998;46:5-18.
25. O'Toole LB. MRC uses checklist similar to CONSORT's. *BMJ*. 1997;314:1127.
26. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF, for the QUOROM group. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. *Lancet*. 1999;354:1896-1900.
27. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA*. 2000;283:2008-2012.
28. Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analysis. *JAMA*. 1996;276:1339-1341.
29. Neumann PJ, Stone PW, Chapman RH, Sandberg EA, Bell CM. The Quality of Reporting in Published Cost-Utility Analyses, 1976-1997. *Ann Intern Med*. 2000;132:964-972.
30. Altman DG. The scandal of poor medical research. *BMJ*. 1994;308:283-284.

sumption and risk of heart failure among older persons. *JAMA*. 2001;285:1971-1977.
20. Cooper HA, Exner DV, Domanski MJ. Light-to-moderate alcohol consumption and prognosis in patients with left ventricular systolic dysfunction. *J Am Coll Cardiol*. 2000;35:1753-1759.
21. Rimm E, Klatsky AL, Grobbee D, Stampfer MJ. Review of moderate alcohol consumption and reduced risk of coronary heart disease. *BMJ*. 1996;312:731-736.
22. Longnecker MP. Alcoholic beverage consumption in relation to risk of breast cancer: meta-analysis and review. *Cancer Causes Control*. 1994;5:73-82.
23. Friedman GD, Klatsky AL. Is alcohol good for your health? [editorial]. *N Engl J Med*. 1993;329:1882-1883.
24. Pearson TA, Terry P. What to advise patients about drinking alcohol. *JAMA*. 1994;272:957-958.

CONSORT Revised—Improving the Reporting of Randomized Trials

Drummond Rennie, MD

IF PHYSICIANS ARE TO BASE TREATMENT DECISIONS ON THE evidence in the medical literature, all the relevant results of trials must be available easily and consistently. Yet it is common to have trouble identifying the hypothesis, the research question, and the design of a published trial. It is even more common to lose count of the participants or to be unable to tell who received what therapies and the type of analysis used. As a result, it is often impossible to know whether the conclusions are justified by the data.

In February 1995, Schulz and colleagues[1] published an important article that drew attention to this sad state of affairs and to the importance of complete reporting of clinical trials if bias was to be avoided. At that time 2 groups, responding to the widespread perception that reporting of study results was highly variable, had held meetings to try to articulate the standards of good reporting. The first, the Standards of Reporting Trials (SORT) group, met in Ottawa, Ontario, in 1993 and published its recommendations in December 1994.[2] The second group met at Asilomar, Calif, in 1994 and also published its recommendations for reporting of clinical trials at the end of that year.[3]

In 1995, *JAMA* editors persuaded the authors of a clinical trial already accepted for publication to rewrite their article to conform with all the SORT recommendations.[4,5] The report turned out to be an extraordinary, one-of-a-kind prototype. The article was structured into more than 30 parts, was apparently agonizing to write, and was certainly a torture to read. But the experiment was instructive because the authors, and numerous readers, let us know their reactions and suggestions. It was apparent that the SORT recommendations were too inflexible, too mechanistic, and too little concerned with the external validity or applicability of the trial results.[5]

Under the leadership of David Moher, the 2 groups then pooled their results.[5] In 1996, the result of this cooperation—the Consolidated Standards of Reporting Trials (CONSORT) statement—was published.[6,7] The statement described the process, provided the rationale for the various reporting requirements listed, and, crucially, allowed a mechanism that permitted considerable individual freedom on the part of journal editors in the way they set forth the information. In addition, CONSORT recommended a flow diagram so that the reader might easily follow the progress of participants through the various stages of the trial, and in every arm of the study.

Some trialists confused CONSORT with the earlier SORT and so damned CONSORT for the very reasons that had brought about the changes found in CONSORT.[8,9] They objected to what they perceived as unilateral imposition of excessively prescriptive rules by nontrialists.[8] This view ignored all the experienced trialists who had taken part in the lengthy process of feedback, of trial and error, that had informed the final document.[9] Fortunately, no substantive criticisms were raised[9] and the general reaction was strongly favorable.

CONSORT was republished in and endorsed by many journals, in several languages (http://www.consort-statement.org). Its use is recommended by the International Committee of Medical Journal Editors, the Council of Science Editors, and the World Association of Medical Editors. CONSORT has been so successful that similar groups of scientists and editors have set up standards, using a similar template and based on empirical evidence whenever possible, to increase the quality of reporting of meta-analysis of randomized trials (QUOROM),[10] and the reporting of the meta-analysis of observational studies (MOOSE).[11] An effort led by Jeroen Lijmer to establish standards to improve the reporting of studies assessing diagnostic tests (STARD) is proceeding, and this group is expected to publish its recommendations soon. Matthias Egger and colleagues are setting up similar groups for improving the reporting of case-control and cohort studies.

Anyone who has read through many hundreds of randomized controlled trials (RCTs) is immediately struck by the fact that when the authors have used the CONSORT checklist and flow diagram, it takes a fraction of the time to get the essential information necessary to assess the quality of a trial. In this issue of THE JOURNAL, Moher et al[12] show in 3 large medical journals, including *JAMA*, that adopted the CONSORT statement, study reporting was improved more than in one such journal that has not chosen to do so. However, before concluding that CONSORT is some sort of panacea, Moher et al also demonstrate that even in journals that have made a strong commitment to CONSORT, reporting was deficient

See also pp 1987, 1992, 1996, and 2000.

Author Affiliation: Dr Rennie is Deputy Editor, *JAMA*.
Disclosure: Dr Rennie has been a member of the Asilomar, CONSORT, QUOROM, MOOSE, and STARD groups.
Corresponding Author and Reprints: Drummond Rennie, MD, *JAMA*, 515 N State St, Chicago, IL 60610 (e-mail: rennie@itsa.ucsf.edu).

©2001 American Medical Association. All rights reserved.

in many ways, for example, in detailing concealment of allocation.[12] This shows that editors have as much difficulty as trialists in learning new behaviors.

But the CONSORT statement, developed after a great deal of discussion and some experimentation, was never intended to be set in concrete. From the first, it was assumed that CONSORT would have to change as new evidence accumulated on the importance of various items of reporting. In 1998, Meinert[13] published a detailed and cogent critique of CONSORT and made several important suggestions for its improvement. In that same issue of *JAMA*, Moher detailed the often woeful ways in which trials were still being reported and stressed the importance of evidence in improving what he considered to be "an evolving tool," CONSORT.[14]

Also in this issue of THE JOURNAL is an article by Egger et al,[15] who examine the use of flow diagrams recommended in the CONSORT statement to show the path of participants from enrollment to analysis. Egger et al found that use of the diagrams was associated with more complete reporting and recommend that all reports of RCTs include them. However, the authors noted problems with the use of the recommended flow diagram, with few of the reports they studied including the number of participants who received treatments as allocated. Moreover, as Meinert pointed out,[13] there were deficiencies in the diagram itself, such as failure to recommend that the number of participants included in the main analysis be specified, a number needed to assess whether an intention-to-treat analysis had been performed. Egger et al make some sensible suggestions for revision of the flow diagram[15] and these revisions have been incorporated in the revised CONSORT statement.

Publication of the revised CONSORT statement in this issue of THE JOURNAL[16] coincides with its simultaneous publication in the *Annals of Internal Medicine* and *The Lancet*, 2 journals that have, in addition to *JAMA*, strongly supported CONSORT. Just as the original SORT and Asilomar statements were revised in response to public comments and experience, so the CONSORT statement, flow diagram, and the checklist have all been revised. The checklist, as before, consists of items that empirical evidence has shown must be included if reporting bias is to be minimized. The changes are listed in the new statement.[16] To help with the use and dissemination of the revised statement, an explanatory and elaboration article is being published with the CONSORT statement in the *Annals of Internal Medicine*.[17] The revision of CONSORT, which is clearer and more flexible, should make it even easier for authors and editors to use and should greatly improve the transparency of reporting. And since the reporting of a trial is inseparable from the rest of its conduct, I expect these revised standards eventually to serve a valuable educational function and improve the way trials are conducted.[18]

Finally, in this issue of THE JOURNAL Devereaux and colleagues[19] report that there is a long way to go to achieve an acceptable degree of precision in thinking about as well as in reporting clinical trials. The authors assessed whom physicians understood to have been blinded (masked) when a study was reported as using single, double, and triple blinding. They found that physicians' interpretations were quite variable, with the respondents offering 10, 17, and 15 unique interpretations, respectively, of the 3 sorts of blinding. When the authors looked at recently published textbooks, they found 5, 9, and 7 different interpretations of each. Lewis Carroll's Humpty Dumpty could say: "When I use a word . . . it means just what I choose it to mean—neither more nor less,"[20] but in the interpretation of science, there is no place for such ambiguity. As Devereaux et al suggest, the answer must be for authors to describe exactly and completely what they did, which is what CONSORT and these other initiatives are all about.

The whole of medicine depends on the transparent reporting of clinical trials. There is plenty of evidence for biased reporting due to commercial influences.[21] To retain credibility, trialists, other researchers, and editors have to show the profession and the public evidence that we are making an earnest attempt to achieve the very highest standards of transparent reporting.

REFERENCES

1. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. *JAMA*. 1995;273:408-412.
2. The Standards of Reporting Trials Group. A proposal for structured reporting of randomized controlled trials. *JAMA*. 1994;272:1926-1931.
3. Working Group on Recommendations for Reporting Clinical Trials in the Biomedical Literature. Call for comments on a proposal to improve reporting of clinical trials in the biomedical literature. *Ann Intern Med*. 1994;121:894-895.
4. Williams JW, Holleman DR, Samsa GP, Simel DL. Randomized controlled trial of 3 versus 10 days of trimethoprim/sulfamethoxazole for acute maxillary sinusitis. *JAMA*. 1995;273:1015-1021.
5. Rennie D. Reporting randomized controlled trials: an experiment and a call for responses from readers. *JAMA*. 1995;273:1054-1055.
6. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA*. 1996;276:637-639.
7. Rennie D. How to report randomized controlled trials: the CONSORT statement. *JAMA*. 1996;276:649.
8. Meade TW, Wald N, Collins R. CONSORT statement on the reporting standards in clinical trials. *BMJ*. 1997;314:1126.
9. Altman DG, Moher D, Rennie D. CONSORT statement on the reporting standards of clinical trials. *BMJ*. 1997;314:1127.
10. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF, for the QUOROM Group. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet*. 1999;354:1896-1900.
11. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA*. 2000;283:2008-2012.
12. Moher D, Jones A, Lepage L, for the CONSORT Group. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA*. 2001;285:1992-1995.
13. Meinert CL. Beyond CONSORT. *JAMA*. 1998;279:1487-1489.
14. Moher D. CONSORT: an evolving tool to help improve the quality of reports of randomized controlled trials. *JAMA*. 1998;279:1489-1491.
15. Egger M, Jüni P, Bartlett C, for the CONSORT Group. Value of flow diagrams in reports of randomized controlled trials. *JAMA*. 2001;285:1996-1999.
16. Moher D, Schulz KF, Altman D, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*. 2001;285:1987-1991.
17. Altman DG, Schulz KF, Moher D, et al, for the CONSORT Group. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001;134:663-694.
18. O'Toole LB. CONSORT statement on the reporting standards of clinical trials: MRC uses checklist similar to CONSORT's. *BMJ*. 1997;314:1127.
19. Devereaux PJ, Manns BJ, Ghali WA, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA*. 2001;285:2000-2003.
20. Carroll L. *Through the Looking-Glass and What Alice Found There*. New York, NY: William Morrow & Co; 1887:chap 6.
21. Rennie D. Fair conduct and fair reporting of clinical trials. *JAMA*. 1999;282:1766-1768.

©2001 American Medical Association. All rights reserved.

SPECIAL ARTICLE

DOES INCREASED ACCESS TO PRIMARY CARE REDUCE HOSPITAL READMISSIONS?

MORRIS WEINBERGER, PH.D., EUGENE Z. ODDONE, M.D., M.H.SC., AND WILLIAM G. HENDERSON, PH.D., FOR THE VETERANS AFFAIRS COOPERATIVE STUDY GROUP ON PRIMARY CARE AND HOSPITAL READMISSION*

Abstract *Background.* For chronically ill patients, readmission to the hospital can be frequent and costly. We studied the effect of an intervention designed to increase access to primary care after discharge from the hospital, with the goals of reducing readmissions and emergency department visits and increasing patients' quality of life and satisfaction with care.

Methods. In a multicenter randomized, controlled trial at nine Veterans Affairs Medical Centers, we randomly assigned 1396 veterans hospitalized with diabetes, chronic obstructive pulmonary disease, or congestive heart failure to receive either usual care or an intensive primary care intervention. The intervention involved close follow-up by a nurse and a primary care physician, beginning before discharge and continuing for the next six months.

Results. The patients were severely ill. Half of those with congestive heart failure (504 patients) had disease in New York Heart Association class III or IV; 30 percent of those with diabetes (751 patients) had end-organ damage; and a quarter of those with chronic obstructive pulmonary disease (583 patients) required home oxygen treatment or oral corticosteroids. The patients had extremely poor quality-of-life scores. Although they received more intensive primary care than the controls, the patients in the intervention group had significantly higher rates of readmission (0.19 vs. 0.14 per month, $P=0.005$) and more days of rehospitalization (10.2 vs. 8.8, $P=0.041$). The patients in the intervention group were more satisfied with their care ($P<0.001$), but there was no difference between the study groups in quality-of-life scores, which remained very low ($P=0.53$).

Conclusions. For veterans discharged from Veterans Affairs hospitals, the primary care intervention we studied increased rather than decreased the rate of rehospitalization, although patients in the intervention group were more satisfied with their care. (N Engl J Med 1996;334: 1441-7.)

©1996, Massachusetts Medical Society.

DESPITE strategies such as prospective payment and required approval for hospitalization before admission, costs of inpatient care in 1993 accounted for $327 billion, or 42 percent of national spending for medical care.[1] Readmissions account for up to half of all hospitalizations and 60 percent of hospital costs.[2-5] Besides the expense, readmissions may reflect poor-quality care.[6-9]

There is pressure both to reduce inpatient services and to deliver high-quality care. One efficient strategy would be to identify patients who are at increased risk for hospital readmission and to provide them with intensive primary care.[8,10] Because of their multiple coexisting medical illnesses, poor functional status, and low socioeconomic status,[11-14] veterans discharged from the General Medicine Service of Veterans Affairs Medical Centers are one such group.[6,15]

We conducted a multicenter randomized, controlled trial of a program designed to increase access to primary care for such veterans. Our primary hypothesis was that this program would reduce the patients' rates of readmission and days of hospitalization during the six months after discharge. We also examined the effects of the program on the time to the first readmission, the proportion of patients readmitted, and the number of emergency department visits, and on health-related quality of life and satisfaction with care.

From the Richard L. Roudebush Veterans Affairs Medical Center (VAMC), Indianapolis (M.W.); the Center for Health Services Research in Primary Care, Durham VAMC, Durham, N.C. (E.Z.O.); and the Center for Cooperative Studies in Health Services, Hines VAMC, Hines, Ill. (W.G.H.). Address reprint requests to Dr. Weinberger at Health Services Research and Development (11H), Richard L. Roudebush VAMC, 1481 W. Tenth St., Indianapolis, IN 46202.

Supported by the Health Services Research and Development Service, Department of Veterans Affairs, through Veterans Affairs Cooperative Study in Health Services No. 8 and the Career Development Program.

*Other contributing authors were Anita Giobbie-Hurder (Hines VAMC, Hines, Ill.), Maria Horner (Durham VAMC, Durham, N.C.), Francis Goldstein (William Jennings Bryan Dorn VAMC, Columbia, S.C.), Paulette Ginier (Fresno VAMC, Fresno, Calif.), Charles Mengel (Dwight D. Eisenhower VAMC, Leavenworth, Kans.), James Huey (Cincinnati VAMC, Cincinnati), David A. Asch (Philadelphia VAMC, Philadelphia), Neil Farber (Philadelphia VAMC), Lawrence Loo (Jerry Pettis Memorial VAMC, Loma Linda, Calif.), Edward Mack (Brooklyn VAMC, Brooklyn, N.Y.), David M. Smith (Richard L. Roudebush VAMC, Indianapolis), Denise M. Hynes (Hines VAMC), and John R. Feussner (Durham VAMC). Additional participating institutions and investigators are listed in the Appendix.

METHODS

Study Sites

This multicenter randomized, controlled trial was conducted at nine Veterans Affairs Medical Centers (see Appendix) chosen for diversity of location and academic affiliation; the sites were not selected on the basis of their readmission rates. Before the investigation began, all the study personnel met to review and standardize the study protocol. The study was approved and reviewed annually by the human rights committee of the Hines Veterans Affairs Cooperative Studies Program Coordinating Center. The Research and Human Subjects Committee of each participating Veterans Affairs Medical Center also approved the study.

Criteria for Eligibility

Patients hospitalized in the General Medicine Service were potentially eligible if they had a diagnosis of diabetes mellitus, chronic obstructive pulmonary disease, or congestive heart failure that was documented in the medical record at or before the time of the index admission (and that was not necessarily the reason for that admission). We selected these three diseases because they are prevalent among veterans, because patients with these diseases are commonly readmitted, and because hospital readmissions to treat these diseases might be reduced if primary care physicians provided intervention to outpatients.

Patients were excluded if they were already receiving continuous care at a primary care clinic (for example, in general medicine or geriatrics); if they were receiving dialysis, chemotherapy, or radiation therapy; if they resided in, or planned to be discharged to, a nursing home; if they were admitted only to undergo a procedure; if they were participating in another active study; if they were hospitalized to rule

out cancer or to receive terminal care; if they did not speak English; if they had a score of 5 or less on the Mental Status Questionnaire[16] and had no care giver; if they refused to give informed consent; or if they had no access to a telephone.

Study Design

Research assistants at each site screened all patients admitted to the General Medicine Service. Potentially eligible patients were referred to the study nurse, who determined their eligibility, obtained informed consent, and collected base-line data. The patients were then randomly assigned to receive either customary post-discharge care or the primary care intervention. They were stratified within each study site according to entitlement status (a variable derived from the patient's service-connection status, which can affect access to outpatient care) and index disease (diabetes, congestive heart failure, chronic obstructive pulmonary disease, or more than one of these). The patients' group assignments were made by telephoning the statistical coordinating center.

All the patients were followed for six months. The research assistant, who was unaware of the patients' group assignments, telephoned the patients 30 and 180 days after randomization to assess their quality of life, satisfaction with care, and use of health care services outside the Veterans Affairs Medical Centers.

Intervention Group

At each site, the intervention was delivered by a team consisting of one licensed registered nurse and one primary care physician. The study nurses were experienced in patient care (mean length of experience, 5.8 years; range, 2 to 10); four had prior clinical experience with the Department of Veterans Affairs. There were 96 attending physicians, 6 fellows in general medicine, and 12 house staff. The attending physicians were predominantly board-certified in internal medicine (70 physicians) or family practice (3 physicians); they had completed medical school a mean of 12.7 years earlier and had a mean of 4.8 years of Veterans Affairs experience.

The intervention (the details of which are specified in Table 1) had both an inpatient component, which began immediately after randomization, and an outpatient component, which began at discharge. When patients assigned to the intervention group were readmitted to the hospital, the inpatient protocol was repeated.

Control Group

We neither required nor prohibited any post-discharge care for the patients in the control group. Their care after discharge could be provided by community physicians or at Veterans Affairs clinics, as arranged by the physicians treating them as inpatients. The control patients did not have access to the primary care nurse and received no supplemental education or assessment of needs beyond what was customarily offered at each site.

Primary and Secondary Outcomes

For outcomes involving the use of Veterans Affairs health services, data were abstracted from the Patient Treatment File (a national administrative data base containing information on all Veterans Affairs hospitalizations) and computer systems at local hospitals (to obtain data on outpatient visits) for 180 days after randomization. The use of non–Veterans Affairs health services was estimated on the basis of reports by the patients. With the patients' permission, we asked non–Veterans Affairs providers identified by the patients to send records of the use of their services during the study period. Only use of services that could be verified by the provider was counted. We computed the total number of days of rehospitalization and rates of readmission per patient, the time to the first readmission, the proportion of patients who were readmitted, the number of emergency department visits, and the number of outpatient visits during the 180-day study period.

Quality of life was measured on the "short form" of a questionnaire (the SF-36) that contains 36 items and has been well validated[17-21] and widely used among veterans.[12-14] Eight scores (physical functioning, physical role functioning, emotional role functioning, social functioning, bodily pain, mental health, vitality, and general perceptions of health) are calculated from the responses, ranging from 0 (for poorest) to 100 (for best); differences of 3 to 5 points are considered important.[18,21] To measure patients' satisfaction with their care, we used the following 11 scales from the Patient Satisfaction Questionnaire[22] that are relevant to veterans: satisfaction with emergency care, convenience of care, access to care, continuity of care, competence, risks (i.e., those at which doctors place their patients), quality of doctor's facilities, expenses, explanations of care, degree of consideration shown, and overall satisfaction. Each scale is scored from 1 (for least satisfied) to 5 (for most satisfied).

Table 1. Components of the Primary Care Intervention and Mean Ratings of Compliance.

Component	Mean Compliance (%)*
Before discharge	
Within three days before discharge, primary care nurse assessed the patient's post-discharge needs, developed a list of medical problems, provided educational materials, assigned the patient to a primary care physician, and gave the patient a card with the names and beeper numbers of the primary care nurse and primary care physician.	99.7
Primary care physician visited the patient personally within two days before discharge to review the hospital course, discharge plans, lists of problems, and medication regimens, discussing discharge plans with hospital physicians as necessary.	74.5
Primary care nurse made an appointment for the patient to visit the primary care clinic within one week of discharge.	62.5
After discharge	
Primary care nurse telephoned the patient within two working days after discharge to assess potential difficulties with medications or medical regimens, identify health problems arising since discharge, make sure that patient knew how to contact providers, and remind patient of the follow-up appointment.	87.3
Patient kept first post-discharge appointment.	82.0
Primary care physician and primary care nurse reviewed and updated the treatment plans at the first post-discharge appointment.	94.1
Appointment reminder sent, if necessary.	50.0
Missed-visit protocol implemented, if necessary.	43.1
Overall score	89.0

*Scores are means for all patients in the intervention group.

Other Variables

Because prior hospitalization is a predictor of future use, we used the Patient Treatment File to extract data on hospital use for all study patients during the 180 days before randomization. According to a validated measure of coexisting conditions,[23] we classified the patients as being at low, medium, or high risk for readmission at base line. The severity of each of the three diseases being studied was estimated by a review of the patient's chart at base line.

Each readmission to a Veterans Affairs Medical Center during the 180-day study period was classified in two ways by a panel of physicians comprising all the principal investigators at each site. First, preventability was assessed on the basis of patient-related, clinical, or systemwide factors that could have averted the readmission[24] (rating forms are available on request). Each readmission was also classified as elective or nonelective. Elective readmissions were those occurring for nonurgent, scheduled procedures (for example, cataract surgery, cardiac catheterization, bronchoscopy, hernia repair, and colonoscopic examination). All other readmissions were considered nonelective.[25]

Intensity of Primary Care

For each patient in the intervention group, each component of the intervention (Table 1) was rated according to whether it had been carried out according to the protocol. Scores for compliance were aggregated for all components to calculate both overall and site-specific scores. Other indicators of the intensity of primary care delivery in-

cluded the number of days from the patient's discharge after the index hospitalization to the first visit to the primary care clinic, the number of visits to that clinic, and the number and duration of telephone calls between patients and primary care nurses during the study period.

Statistical Analysis

Assuming a two-sided significance level of 0.05 and 85 percent power, we calculated that 700 patients would be needed in each group in order to detect 28 percent reductions in both primary outcomes (readmission rates and days of rehospitalization).

The success of randomization was determined by comparing the base-line characteristics of the intervention group with those of the control group by the chi-square test (for categorical variables), Student's t-test (for continuous variables), and the Wilcoxon rank-sum test (for non-normally distributed variables).

Patients continued in the study for 180 days unless they died or withdrew. For those who did not remain in the study for the entire 180 days, the data collected between randomization and the date of censoring were used in all the analyses (that is, we conducted an intention-to-treat analysis). Our main analyses used Wilcoxon rank-sum tests (of readmission rates, days of rehospitalization, emergency department visits, and outpatient visits), chi-square tests (of the proportion of patients readmitted), and Kaplan–Meier estimates of survival and log-rank tests (of the time to readmission). Because the measures of quality of life and patients' satisfaction were multidimensional, we studied them by multivariate analysis of variance. Finally, we performed an analysis of covariance for our primary and secondary outcomes. The covariates included the stratification variables and the number of hospital days during the 180 days before randomization.

Results

Study Patients

From November 1992 through July 1994, 10,129 patients were screened; 3209 met all eligibility criteria, and 1396 of them (43.5 percent) were randomized (range per site, 116 to 202 patients). The most common reasons for the nonenrollment of eligible patients were the patient's decision not to participate (971 patients) and discharge from the hospital before randomization (446 patients). Patients who declined to participate typically did so because they had an established relationship with a specialist or a non–Veterans Affairs physician and did not wish to risk random assignment to the care of a new physician.

The demographic and clinical characteristics of the patients are shown in Table 2. The patients had a substantial burden of illness at base line: half of those with congestive heart failure had disease in New York Heart Association class III or IV; one third of the diabetic patients had objective evidence of end-organ damage; and one quarter of the patients with chronic obstructive pulmonary disease required oxygen treatment at home, oral corticosteroids, or both. Moreover, two thirds of the study patients had risk scores that placed them at medium or high risk for readmission. There were no statistically significant differences between the study groups, although there was a trend toward more hospital days during the 180 days before randomization in the intervention group (4.6 days, vs. 3.9 days in the control group; P = 0.09).

The patients' extensive burden of illness was also evident from their extremely poor base-line scores for quality of life (Table 3). The patients were moderately satisfied with most aspects of their care, giving it an average score of approximately 3 on the 5-point scales for satisfaction (Table 3). There were no base-line differences between the study groups in quality of life or satisfaction with care. During the study period, 106 patients died (59 in the intervention group and 47 in the control group), and 16 withdrew their consent to be studied (11 and 5, respectively). There were no significant differences between the study groups in the follow-up status of the patients (P = 0.13).

Table 2. Base-Line Characteristics of the Patients According to Study Group.*

Characteristic	Intervention Group (N = 695)	Control Group (N = 701)
Age (yr)	63.0±11.1	62.6±10.9
Education (yr)	11.2±3.2	11.0±3.2
Marital status (% married)	52.4	54.6
Race (%)		
White, non-Hispanic	64.2	65.9
Black, non-Hispanic	28.5	26.4
Other	7.3	7.7
Male sex (%)	99.0	98.0
Any employment, full- or part-time (%)	17.7	16.4
Length of index admission (days)	10.1±14.3	10.4±13.1
Time from randomization to discharge (days)	6.6±13.0	6.8±11.6
Eligibility status (% service-connected)	28.6	29.7
Hospital days during the 180 days before randomization	4.6±10.6	3.9±10.4
Index diagnosis (% of patients)		
Diabetes mellitus	35.7	35.0
Congestive heart failure	13.2	13.4
Chronic obstructive pulmonary disease	23.2	23.4
>1 diagnosis	27.9	28.2
Disease status†		
Diabetes (no. of patients)	376	375
Not using insulin, no end-organ disease (%)	45.0	42.9
Using insulin, no end-organ disease (%)	26.9	24.5
End-organ disease (%)	28.2	32.5
Congestive heart failure (no. of patients)‡	249	255
NYHA class I (%)	12.9	11.8
NYHA class II (%)	38.1	36.1
NYHA class III (%)	32.9	32.9
NYHA class IV (%)	16.1	19.2
Chronic obstructive pulmonary disease (no. of patients)	295	288
No home oxygen or corticosteroid use (%)	74.9	75.7
Home oxygen, corticosteroid use, or both (%)	25.1	24.3
Risk of readmission (% of patients)		
Low	35.7	34.4
Medium	38.1	42.6
High	26.2	23.0

*Plus–minus values are means ±SD. There were no significant base-line differences between groups.

†The total number of patients shown in the three disease categories exceeds the total number of patients in the study (1396) because some patients had more than one disease.

‡NYHA denotes New York Heart Association.

Intensity of Primary Care

Table 1 shows data on the patients' compliance with the intervention protocol. The mean composite score for compliance was 89 percent (range among sites, 83 to 93 percent). The median time from discharge after the index hospitalization to the first visit to a general medicine clinic was significantly shorter in the intervention group (7 vs. 13 days, P<0.001). The patients in the intervention group were more likely than the controls

to visit at least one general medicine clinic during the study period (93 percent vs. 77 percent, P<0.001). Over the six-month study period, the patients in the intervention group also made 68 percent more visits to general medicine clinics (mean, 3.7 vs. 2.2; P<0.001) and 5 percent fewer visits to subspecialty clinics (2.1 vs. 2.2, P=0.010) than did the control patients. Otherwise, use of outpatient services, including visits to the emergency department (1.9 vs. 1.7, P=0.12), was similar in the two groups. In addition to providing care during visits, primary care nurses talked with the patients in the intervention group by telephone a mean of 7.5 times during the study period (range, 4.0 to 13.4 calls), for an average of 5.7 minutes per call. Use of outpatient services outside Veterans Affairs Medical Centers was infrequent in both groups (intervention group, 7 percent; controls, 9 percent; P=0.55).

Hospital Use

Although the patients in the intervention group had contact with the primary care team, the effect of the intervention on hospital use was contrary to that predicted by our hypothesis (Table 4). The intervention group had a higher monthly readmission rate than the control group (0.19 vs. 0.14 readmission, P=0.005) and more days of rehospitalization (10.2 vs. 8.8, P=0.041). Approximately 5 percent of the use of inpatient services by both study groups occurred in non–Veterans Affairs hospitals. We also found trends suggesting that a higher proportion of the patients in the intervention group were readmitted (49 percent vs. 44 percent, P=0.06), and were readmitted sooner (P=0.07), than was the case in the control group. Finally, more patients in the intervention group (P=0.052) had multiple readmissions during the study period (Fig. 1).

Table 3. Base-Line Scores on Questionnaires Measuring Quality of Life and Satisfaction with Care, According to Study Group.*

Measure	Intervention Group (N=695)	Control Group (N=701)
	mean ±SD	
Quality of life (SF-36)		
Physical functioning	44.4±29.9	43.6±30.2
Physical role functioning	23.0±37.1	20.8±34.5
Emotional role functioning	55.6±45.8	56.1±45.4
Social functioning	55.9±32.0	54.5±32.6
Bodily pain	50.5±33.0	48.5±31.9
Mental health	65.5±24.0	64.1±23.4
Vitality	33.7±25.5	31.1±24.4
General perceptions of health	42.2±21.9	39.4±21.3
Patient Satisfaction Questionnaire		
Nonfinancial access to care		
Emergency care	3.3±0.7	3.4±0.7
Convenience	3.5±0.8	3.4±0.9
Access to care	3.2±0.9	3.1±0.9
Continuity of care	2.7±1.0	2.7±1.0
Technical quality of care		
Quality — competence	3.4±0.5	3.4±0.5
Prudence — risks	3.0±0.5	3.1±0.5
Doctor's facilities	3.6±0.7	3.6±0.7
Prudence — expenses	3.1±0.7	3.1±0.7
Interpersonal manner		
Explanations	3.1±0.9	3.1±0.8
Consideration	3.6±0.5	3.6±0.6
Overall satisfaction	3.2±0.7	3.2±0.7

*Scores on the SF-36 questionnaire for the patients' quality of life range from 0 (lowest) to 100 (highest). On the Patient Satisfaction Questionnaire, scores range from 1 (least satisfied) to 5 (most satisfied). There were no statistically significant differences between groups.

Table 4. Effect of the Intervention on the Use of Inpatient Services, According to Diagnosis and Study Group.*

Variable	Intervention Group (N=695)	Control Group (N=701)	P Value
No. of readmissions/mo			
All patients	0.19±0.4	0.14±0.2	0.005
Diabetes	0.13±0.2	0.11±0.2	
CHF	0.27±0.7	0.15±0.3	
COPD	0.19±0.3	0.14±0.2	
>1 diagnosis	0.23±0.3	0.17±0.2	
Days of rehospitalization			
All patients	10.2±19.8	8.8±19.7	0.041
Diabetes	7.8±18.1	8.0±21.4	
CHF	11.7±26.6	6.8±14.8	
COPD	11.7±20.2	8.3±15.5	
>1 diagnosis	11.4±17.5	11.1±22.4	
Proportion readmitted (%)			
All patients	49.4	44.2	0.06
Diabetes	37.1	38.0	
CHF	52.2	41.5	
COPD	53.4	48.2	
>1 diagnosis	60.3	50.0	

*Plus–minus values are means ±SD. Because the distributions of monthly readmission rates and days of rehospitalization were non-normal, Wilcoxon rank-sum tests were used to calculate P values. CHF denotes congestive heart failure, and COPD chronic obstructive pulmonary disease.

Although the study lacked adequate power to permit us to conduct subgroup analyses according to disease category, greater hospital use was observed in the intervention group in all three disease strata. There were no significant differences between groups with regard to either nonelective (80 percent vs. 77 percent, P=0.22) or preventable (35 percent vs. 37 percent, P=0.57) readmissions. After we adjusted for the stratification variables and the number of days spent in the hospital during the six months before randomization, the difference between the groups in monthly readmission rates remained significant (0.21 vs. 0.15, P<0.001), although the difference in the number of days of rehospitalization did not (10.5 vs. 9.2, P=0.23).

Quality of Life and Satisfaction with Care

Follow-up interviews were completed with 87 percent and 83 percent of the patients available for interview at one and six months, respectively. In both groups, patients' scores on the SF-36 questionnaire were low at base line and remained so throughout the study period, with no significant differences between groups at one month (P=0.99) or six months (P=0.53).

The patients in the intervention group were significantly more satisfied (P<0.001) with their care than the controls at one month (data not shown), and this difference persisted at six months (Fig. 2). Although the group differences were consistent on most of the 11 scales of the Patient Satisfaction Questionnaire that we used, the differences were greatest with regard to the

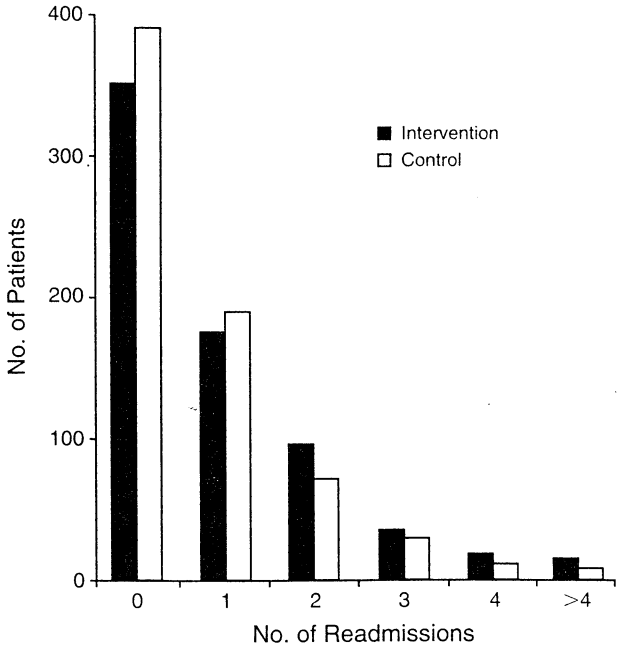


Figure 1. Effect of Intervention on the Number of Readmissions.
P = 0.052 for the comparison between the study groups, by the chi-square test.

patients' perceptions of the continuity of their care (33 percent) and "nonfinancial access to medical care" (i.e., barriers other than financial ones) (16 percent).

Discussion

We examined the effect of an intervention designed to enhance primary care for medically vulnerable patients who were discharged from the General Medicine Service at nine Veterans Affairs Medical Centers. The intervention was designed to improve their access to primary care providers, the coordination of outpatient services, and the provision of comprehensive and continuous care.[26] Although the intensity of primary care was successfully increased, patients receiving this intervention used the hospital significantly more during the six-month study period. At only one of the nine sites were readmissions reduced, and that site was at or below the median level of compliance with the intervention protocol, the frequency of primary care visits per patient, and the extent of telephone contact between patients and nurses. The primary care intervention did not affect the quality of life of the patients who received it, but they were substantially more satisfied than the controls with their care.

What may account for these findings? First, the premise that comprehensive primary care may reduce the use of inpatient services by vulnerable patients may be wrong, at least in the short term. The patients in this study had major complications of their chronic diseases and poor quality of life. They were also at higher risk for readmission than vulnerable patients discharged from a municipal teaching hospital.[27,28] The primary care offered to these seriously ill patients may have led to the detection and treatment of previously undetected medical problems. Second, greater access to primary care providers could have improved communication and, in turn, increased readmissions. Having a channel to voice their complaints can lead to more readmissions among severely ill patients. With a longer period of follow-up, the patients and primary care teams may become more accustomed to each other, and perhaps readmissions would be diminished over time. Finally, the patients in the intervention group may have been sicker than the controls at base line. There was a trend toward more use of inpatient services in this group than among the controls during the six months before enrollment, but even when we adjusted for this prior use there was still a significant difference in readmission rates, although the difference in the number of days of rehospitalization was no longer significant.

Did the intervention harm patients? The patients receiving it were readmitted more frequently and spent more time in the hospital than those receiving usual care, but there is no evidence that the patients receiving the intervention experienced decreased quality of life. Furthermore, at one and six months they were substantially more satisfied with their care.

How do our findings compare with those of earlier randomized trials of interventions designed to reduce hospital readmissions? Facilitating linkages between in-

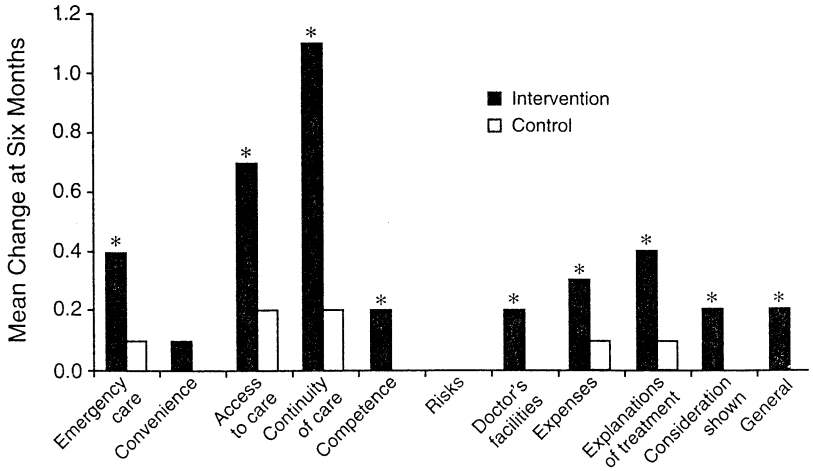


Figure 2. Effect of Intervention on Patients' Satisfaction with Their Care.
Bars show the difference between the six-month scores and the base-line scores. P<0.001 for the comparison between groups by multivariate analysis of variance at both one and six months. Asterisks indicate dimensions for which the difference in the mean change (on a five-point scale) between the study groups was statistically significant (P<0.001). Subscales for which no bar is shown (e.g., Risks) had a mean change of zero.

patient and outpatient care at an urban teaching hospital reduced both the use of services after discharge[27] and the costs of health care[28] among the patients at highest risk for readmission. However, providing access to a case manager at discharge from a Veterans Affairs Medical Center and thereafter had no effect on readmissions.[29] Intensive discharge planning may reduce hospital readmissions in the short term, but not in the long term.[7,30] A single-site study of a multidisciplinary intervention directed by an experienced cardiovascular nurse reduced the use of inpatient services by high-risk elderly patients with congestive heart failure during the 90 days after discharge.[31] This success may be attributed to the use of an intensive, disease-specific protocol that included individualized dietary assessment and instruction, analysis of medications by a geriatric cardiologist, and home visits after discharge.

This study has several limitations. First, it was conducted largely among disadvantaged men receiving care at Veterans Affairs Medical Centers. The generalizability of our findings may also be affected by the presence of systematic differences between the study patients and the eligible patients who did not enroll. However, the 971 eligible patients who declined to participate in the study did not differ from the study patients during the six months after screening with regard to rates of readmission to Veterans Affairs Medical Centers, the number of days of rehospitalization, or the proportion of patients readmitted (data not shown). Second, all the patients were screened while they were hospitalized. Intensive primary care interventions designed for outpatients who were not identified in the hospital might have produced different results. Third, we cannot assess the incremental value of specialty care among patients receiving primary care. For such patients, the mix of generalist and specialist care may be as important as the coordination of care by a primary care team. Finally, the intervention may have improved the quality of care despite the increased rate of readmission.

Our findings show that tremendous resources were needed to sustain the health of this vulnerable group of veterans. If these findings also apply to other groups of medically and socioeconomically disadvantaged patients (for example, the Medicaid population), health care administrators may be reluctant to provide care unless a high capitation rate is established. In addition, the patients in the intervention group were significantly more satisfied with their care than the controls. This difference in satisfaction exceeds that previously shown to result in patients' changing their health care providers.[32-34] Such information is critical for the design of health care systems, given that patients base their decisions about where they will obtain their care largely on subjective ratings of their own experiences. Finally, and perhaps most important, this study highlights the need to evaluate assumptions about changes in health care systems rigorously, across the entire spectrum of patients and diseases. We need a better understanding of how to optimize the care of vulnerable patients, so that the use of services is reduced and the quality of care enhanced.

Appendix

The following additional persons and institutions participated in this investigation: *Chairperson's Office* — P. Landsman and M. Monger (Durham, N.C.). *Hines Center for Cooperative Studies in Health Services* — D. Cavello and R. Lott (Hines, Ill.). *Participating Veterans Affairs Medical Centers* — E. Anteola and R. Varano (Brooklyn, N.Y.); V. Hedger and J. Schultz (Cincinnati); G. Allen and J. Calkins (Columbia, S.C.); A. Ward and M. Foy (Durham, N.C.); P. Hensley and K. Cox (Fresno, Calif.); G. Redmon (Indianapolis); L. Carrel and M. Cook (Leavenworth, Kans.); E. Wise and N. Gordon (Loma Linda, Calif.); and A. Cooney and J. Havey (Philadelphia). *Executive Committee* — C. Ashton, T. Adams, J. Demakis, and J. Gibbs. *Data Monitoring Board* — T. Meyer (chairperson), M. Foulkes, M. Hlatky, and K. Nichol. *Human Rights Committee (Hines, Ill.)* — T. Bering, T. Burris, A. Cole, E. Collins, M. D'Arcy, M. Emanuele, Z. Flournoy-Gill, S. Sanders, T. Schmid, A. Henrick, S. Braithwaite, W. Knopp, W. Juneau, and R. Hahn. *Veterans Affairs Central Office* — D. Deykin and J. Gold (Boston); and J. Gough, S. Meehan, C. Smith, and C. Welch (Washington, D.C.).

References

1. Medicare and Medicaid statistical supplement, 1995. Health Care Financ Rev 1995;16(Stat Suppl/Fall):159.
2. Zook CJ, Savickis SF, Moore FD. Repeated hospitalization for the same disease: a multiplier of national health costs. Milbank Q 1980;58:454-71.
3. Zook CJ, Moore FD. High-cost users of medical care. N Engl J Med 1980; 302:996-1002.
4. Anderson GF, Steinberg EP. Hospital readmissions in the Medicare population. N Engl J Med 1984;311:1349-53.
5. *Idem.* Predicting hospital readmissions in the Medicare population. Inquiry 1985;22:251-8.
6. Holloway JJ, Medendorp SV, Browning J. Risk factors for early readmission among veterans. Health Serv Res 1990;25:213-37.
7. Evans RL, Hendricks RD. Evaluating hospital discharge planning: a randomized clinical trial. Med Care 1993;31:358-70.
8. Weinberger M, Oddone E. Strategies to reduce hospital readmissions: a review. Q Rev Bull 1989;15:255-60.
9. Ashton CM, Kuykendall DH, Johnson ML, Wray NP, Wu L. The association between the quality of inpatient care and early readmission. Ann Intern Med 1995;122:415-21.
10. Safran C, Phillips RS. Interventions to prevent readmission: the constraints of cost and efficacy. Med Care 1989;27:204-11.
11. Waite K, Oddone E, Weinberger M, Samsa G, Foy M, Henderson W. Lack of association between patients' measured burden of disease and risk for hospital readmission. J Clin Epidemiol 1994;47:1229-36.
12. Weinberger M, Kirkman MS, Samsa GP, et al. A nurse-coordinated intervention for primary care patients with non-insulin-dependent diabetes mellitus: impact on glycemic control and health-related quality of life. J Gen Intern Med 1995;10:59-66.
13. Weinberger M, Oddone EZ, Samsa GP, Landsman PB. Are health-related quality of life measures affected by the mode of administration? J Clin Epidemiol 1996;49:135-40.
14. Kazis LE, Clark JA, Miller D, et al. Quality of life in veterans: preliminary results from the Veterans Health Study. In: Submitted abstracts for the 11th Annual HSR&D Service Meeting, Washington, D.C., April 19–21, 1993. Washington, D.C.: Department of Veterans Affairs, 1993:57. abstract.
15. Ludke RL, MacDowell NM, Booth BM, Hunter SA. Appropriateness of admissions and discharges among readmitted patients. Health Serv Res 1990; 25:501-25.
16. Kahn RL, Goldfarb AI, Pollack M, Peck A. Brief objective measures for the determination of mental status in the aged. Am J Psychiatry 1960;117:326-8.
17. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30: 473-83.
18. Stewart AL, Greenfield S, Hays RD, et al. Functional status and well-being of patients with chronic conditions: results from the Medical Outcomes Study. JAMA 1989;262:907-13. [Erratum, JAMA 1989;262:2542.]
19. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-item short-form health survey (SF-36). II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31:247-63.
20. McHorney CA, Ware JE Jr, Rogers W, Raczek AE, Lu JF. The validity and relative precision of MOS short- and long-form health status scales and Dartmouth COOP charts: results from the Medical Outcomes Study. Med Care 1992;30:Suppl:MS253-MS265.

21. Ware JE Jr. SF-36 health survey: manual and interpretation guide. Boston: Health Institute, New England Medical Center, 1993.
22. Ware JE Jr, Snyder MK, Wright WR, Davies AR. Defining and measuring patient satisfaction with medical care. Eval Program Plann 1983;6:247-63.
23. Smith DM, Norton JA, McDonald CJ. Nonelective readmissions of medical patients. J Chronic Dis 1985;38:213-24.
24. Oddone EZ, Weinberger M, Hurder A, Henderson WG, Feussner JR. Classifying general medicine admissions: are they preventable? Clin Res 1993; 41:819A. abstract.
25. Smith DM, Weinberger M, Katz BP. Predicting nonelective hospitalization: a model based on risk factors associated with diabetes mellitus. J Gen Intern Med 1987;2:168-73.
26. Starfield B. Primary care: concept, evaluation, and policy. New York: Oxford University Press, 1992.
27. Smith DM, Weinberger M, Katz BP, Moore PS. Postdischarge care and readmissions. Med Care 1988;26:699-708.
28. Weinberger M, Smith DM, Katz BP, Moore PS. The cost-effectiveness of intensive postdischarge care: a randomized trial. Med Care 1988;26:1092-102.
29. Fitzgerald JF, Smith DM, Martin DK, Freedman JA, Katz BP. A case manager intervention to reduce readmissions. Arch Intern Med 1994;154:1721-9.
30. Naylor M, Brooten D, Jones R, Lavizzo-Mourey R, Mezey M, Pauly M. Comprehensive discharge planning for the hospitalized elderly: a randomized clinical trial. Ann Intern Med 1994;120:999-1006.
31. Rich MW, Beckham V, Wittenberg C, Leven CL, Freedland KE, Carney RM. A multidisciplinary intervention to prevent the readmission of elderly patients with congestive heart failure. N Engl J Med 1995;333:1190-5.
32. Marquis MS, Davies AR, Ware JE Jr. Patient satisfaction and change in medical care provider: a longitudinal study. Med Care 1983;21:821-9.
33. Davies AR, Ware JE Jr. Involving consumers in quality care assessment. Health Aff (Millwood) 1988;7(1):33-48.
34. Ware JE Jr, Davies AR. Behavioral consequences of consumer dissatisfaction with medical care. Eval Program Plann 1983;6:291-7.

©Copyright, 1996, by the Massachusetts Medical Society
Printed in the U.S.A.

Efficacy of 3 Commonly Used Hearing Aid Circuits

A Crossover Trial

Vernon D. Larson, PhD
David W. Williams, MS
William G. Henderson, PhD
Lynn E. Luethke, PhD
Lucille B. Beck, PhD
Douglas Noffsinger, PhD
Richard H. Wilson, PhD
Robert A. Dobie, MD
George B. Haskell, PhD
Gene W. Bratt, PhD
Janet E. Shanks, PhD
Patricia Stelmachowicz, PhD
Gerald A. Studebaker, PhD
Allen E. Boysen, PhD
Amy Donahue, PhD
Rinaldo Canalis, MD
Stephen A. Fausti, PhD
Bruce Z. Rappaport, PhD
for the Participants of the NIDCD/VA Hearing Aid Clinical Trial Group

Context Numerous studies have demonstrated that hearing aids provide significant benefit for a wide range of sensorineural hearing loss, but no carefully controlled, multicenter clinical trials comparing hearing aid efficacy have been conducted.

Objective To compare the benefits provided to patients with sensorineural hearing loss by 3 commonly used hearing aid circuits.

Design Double-blind, 3-period, 3-treatment crossover trial conducted from May 1996 to February 1998.

Setting Eight audiology laboratories at Department of Veterans Affairs medical centers across the United States.

Patients A sample of 360 patients with bilateral sensorineural hearing loss (mean age, 67.2 years; 57% male; 78.6% white).

Intervention Patients were randomly assigned to 1 of 6 sequences of linear peak clipper (PC), compression limiter (CL), and wide dynamic range compressor (WDRC) hearing aid circuits. All patients wore each of the 3 hearing aids, which were installed in identical casements, for 3 months.

Main Outcome Measures Results of tests of speech recognition, sound quality, and subjective hearing aid benefit, administered at baseline and after each 3-month intervention with and without a hearing aid. At the end of the experiment, patients ranked the 3 hearing aid circuits.

Results Each circuit markedly improved speech recognition, with greater improvement observed for soft and conversationally loud speech (all 52-dB and 62-dB conditions, $P \leq .001$). All 3 circuits significantly reduced the frequency of problems encountered in verbal communication. Some test results suggested that CL and WDRC circuits provided a significantly better listening experience than PC circuits in word recognition ($P = .002$), loudness ($P = .003$), overall liking ($P = .001$), aversiveness of environmental sounds ($P = .02$), and distortion ($P = .02$). In the rank-order ratings, patients preferred the CL hearing aid circuits more frequently (41.6%) than the WDRC (29.8%) and the PC (28.6%) ($P = .001$ for CL vs both WDRC and PC).

Conclusions Each circuit provided significant benefit in quiet and noisy listening situations. The CL and WDRC circuits appeared to provide superior benefits compared with the PC, although the differences between them were much less than the differences between the aided vs unaided conditions.

JAMA. 2000;284:1806-1813 www.jama.com

SENSORINEURAL HEARING LOSS IS one of the most prevalent disabling conditions reported in the United States, affecting some 20 million to 26 million people.[1-3] Hearing loss is present in 35% to 42% of individuals older than 65 years.[4-6] It adversely affects physical, cognitive, behavioral, and social function, as well as the general quality of life,[7] and has been linked to depression and dementia.[8-10]

While hearing aids are the main form of treatment, only about 20% of those who could benefit from hearing aids wear them.[2,3,11] Moreover, surveys have suggested that about 50% of hearing aid users are dissatisfied with their instruments.[12] A recent survey, however,

Author Affiliations and a complete list of the members of the NIDCD/VA Hearing Aid Clinical Trial group are listed at the end of this article.
Corresponding Author and Reprints: Vernon D. Larson, PhD, Howard Leight Industries, 7828 Waterville Rd, San Diego, CA 92154 (e-mail: vlarson@howardleight.com).

 ©2000 American Medical Association. All rights reserved.

suggested that because of improved technology, approximately 65% of hearing aid users are satisfied with their devices.[13]

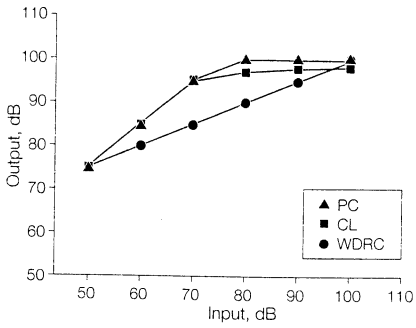
A vast array of hearing aid technologies is available, ranging from simple and relatively inexpensive analog circuits to complex and expensive digital devices that require sophisticated fitting procedures. Whereas numerous studies have demonstrated that hearing aids provide significant benefit compared with unaided listening for persons with hearing losses ranging from mild to severe,[14-17] carefully controlled, multicenter clinical trials of the relative benefit provided by different types of hearing aids have not been conducted. Laboratory studies and small-scale field studies have been designed in ways that make them difficult to compare and have failed to show consistent superiority for any type of signal processing.[18,19]

Choices among available hearing aids must be made without the benefit of data from well-designed clinical trials. This report presents the results of a double-blind, multicenter clinical trial to compare the efficacy of 3 different hearing aid circuits. Efficacy was measured in a variety of listening situations using tests of speech understanding, sound quality, and patient rank-order ratings. The 3 hearing aid circuits jointly account for 70% of the US hearing aid market.

The clinical trial compared 3 commonly used hearing aid circuits: the linear peak clipper (PC), the compression limiter (CL), and the wide dynamic range compressor (WDRC). The PC and CL circuits amplify input sounds linearly up to a predetermined level (usually set relative to loudness discomfort levels). Above that level, the output is limited using 2 different electronic methods. FIGURE 1 illustrates the major difference among the 3 circuits. For the PC, as the input signal level increases by 10 dB, so does the output level up to its maximum output capabilities when the instrument is said to be in "saturation." The CL operates similarly in that the output increases linearly up to a certain point. After that, however, the output is reduced by circuitry that automatically turns down the gain of the hearing aid by a fixed ratio. In this instance, the output is allowed to increase by 1 dB for each 10-dB increase in the input sound level. Finally, the WDRC behaves similarly to the CL circuit except that the automatic gain function begins at lower input sound levels and allows, in this instance, the output to increase 1 dB for each 2 dB increase in input sound level up to its point of maximum output.

The PC removes the positive and/or negative peaks of the amplified signal, whereas the CL uses automatic volume control circuitry. A disadvantage of PC circuitry is that some acoustic distortion results when the output limit is exceeded.[20,21] Far less distortion is created by CL circuitry.[22] The WDRC circuit allows input signals that vary in level over a wide range to be amplified as a narrower range of output signals,[23,24] which is associated with the reduced dynamic range found in the majority of sensorineural hearing loss. Although theoretically beneficial to listener comfort and speech understanding, a disadvantage of compression circuits (eg, CL and WDRC) is that they alter the temporal characteristics of signals in a way that can be apparent to the listener.[25,26]

Figure 1. Performance Characteristics of 3 Hearing Aid Circuits



PC indicates linear peak clipper; CL, compression limiter; and WDRC, wide dynamic range compressor.

METHODS

Trial Design

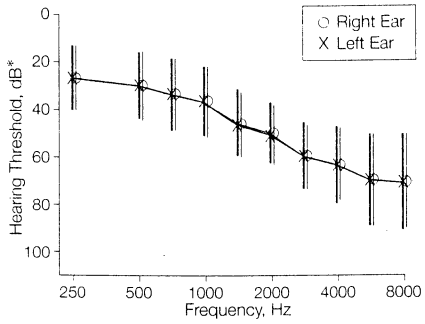
Eight audiology laboratories located within Department of Veterans Affairs (VA) medical centers participated. The experimental design was a 3-period, 3-treatment crossover design. Baseline measurements were made using a battery of tests in the unaided condition (no hearing aids). Patients were then stratified by center and randomized to 1 of 6 sequences of the 3 hearing aid circuits. Six sequences were used so that each hearing aid circuit had approximately an equal number of patients who used the circuit first, second, and third. Each block of 6 consecutive patients within each center was balanced so that each sequence was represented once. The actual frequencies for the 6 sequences in the trial ranged from 59 to 61. In each of the 3 periods, the patients were fit binaurally and used 1 circuit (aided condition) for 3 months. At the end of each period, the battery of outcome tests was repeated in both the unaided and aided conditions. The protocol was conducted in double-blind fashion. Neither the audiologist who administered the tests nor the patient could identify the circuit being worn because all 3 hearing aid circuits resided in the same casement and because a different audiologist programmed the device.

Patient Sample

The protocol was approved by the National Institute on Deafness and Other Communication Disorders (NIDCD) Hearing Aid Advisory Committee, the Hines VA Cooperative Studies Program Coordinating Center's Human Rights Committee, and by the institutional review board of each participating center. All patients provided informed consent, were fluent speakers of English, and had bilaterally symmetrical sensorineural losses with no evidence of retrocochlear pathology. Average audiometric thresholds for 500, 1000, 2000, 3000, and 4000 Hz were no better than 25-dB hearing level in either ear, with no threshold from 500 to 2000 Hz exceeding 70-dB hearing level. FIGURE 2 shows the mean (±1 SD) audiogram. To ensure that the sample included patients who were typical of the majority of adult hearing aid users, monaural word recognition scores

©2000 American Medical Association. All rights reserved.

Figure 2. Mean Audiogram of Patient Sample (N=360)



Error bars for each ear indicate the range from 1 to −1 SD; the darker error bar on the left in each pair corresponds to the left ear. Asterisk indicates values based on American National Standards Institute guidelines for 1996.

on a recorded version of the Central Institute for the Deaf W-22 test[27] were required to be at least 28%, with a difference no greater than 26% between ears.

Experimental Apparatus

Each participant was fit binaurally with single channel, programmable, full-concha in-the-ear hearing aids (Dyna P2, Phonak, Stäfa, Switzerland) that contained all 3 circuit options. The National Institute of Standards and Technology evaluated prototypes to ensure that characteristics of the hearing aid conformed to the manufacturer's specifications.

The 3 programmable options were PC, CL, and WDRC. The CL had an 8:1 compression ratio (above compression threshold, an 8-dB increase in the input level resulted in only a 1-dB increase in the output) and duration-dependent release time capability. The WDRC had a fixed-compression threshold (approximately 52-dB input sound pressure level [SPL]), a compression ratio that ranged from 1.1:1 to 2.7:1 and a short, fixed release time (50 milliseconds). The maximum output levels of the 3 circuits were programmable over approximately the same range of SPLs.

Electroacoustic measurements[28] were made at each visit to ensure that hearing aid characteristics remained stable. Acoustic gain targets were established using the NAL-R method,[29] and probe microphone procedures were used to verify that targets had been achieved. Maximum output targets were obtained using loudness discomfort levels[30] and were subsequently held constant across visits and circuit types.

Outcome Test Battery

All testing was carried out in audiometric test rooms using identical equipment for test presentation and data collection at each site. Three categories of outcome measures were used: speech recognition tests, category ratings of perceived sound quality, and self-assessed subjective ratings of hearing aid benefit.

Speech Recognition

Two tests of speech recognition were used. A recorded version of a monosyllabic word-recognition test, the NU-6,[27] was presented using a single loudspeaker (positioned at 0°; ie, patients faced the loudspeaker) at an SPL of 62 dB (conversational speech level). Each of the 4 NU-6 lists contains 50 scoreable items with each item having a value of 2%. At conversational speech levels, listeners with normal hearing obtain perfect monosyllabic word recognition scores. The second test, a recorded version of the Connected Speech Test (CST),[31,32] consists of 48 passages with 8 to 10 sentences that approximate everyday, connected discourse. Because it was unlikely that a single laboratory condition could represent the range of possible listening conditions, we conducted this test in a variety of presentation and background-noise levels. The CST was presented via the loudspeaker (located at 0° azimuth) at a level of 74-dB SPL (loud speech) in quiet and then again at 74 dB in 3 background noise conditions. For SPLs of 52 dB (soft) and 62 dB (conversational loudness), the speech materials were presented in 3 conditions of background noise.

The background noise used was an uncorrelated multitalker babble,[31] which was delivered from loudspeakers located at azimuths 45° left and right at nominal signal-to-babble (S/B) ratios of −3 dB, 0 dB, and 3 dB. The S/B ratio refers to the relationship of the SPL of the speech to the SPL of the background babble. The nominal 0-dB S/B condition was estimated during the baseline visit prior to conducting tests for each patient by presenting CST practice materials at 62-dB SPL in the "unaided" condition using a bracketing procedure in which the binaural babble level was varied for each subject to produce 50% intelligibility. (The mean [SD] level of the babble was 55 [5.4] dB.) This relationship for each subject was designated as the 0-dB condition. The same S/B ratio for each subject was used for the −3-dB and 3-dB conditions, and for all tests conducted over the 9-month protocol, the same ratios were used. Normal listeners typically receive perfect scores at loud and conversational levels in a quiet background, but their performance at softer levels and in the presence of background noise varies as a function of the difficulty of the listening situation.

Category Ratings

The Quality Rating Test, was used to assess 3 aspects of patients' perception of sound quality: loudness, noise interference, and overall liking of the listening experience. The patients rated each dimension on a 10-point scale. On the loudness scale, 1 was too soft; 10, too loud; and 5, comfortably loud. For overall liking, 1 was very poor or terrible and 10 was excellent. In this task, the patients were instructed to ignore the loudness of the speech and consider only the overall sound quality. For noise interference, they assigned a 1 if noisiness completely interfered with quality and understanding of the speaker and 10 if noisiness did not interfere. Intermediate integer ratings could be assigned for all tests. Sentences designated as practice sentences of the CST[31] served as the stimuli for the Quality Rating Test. Patients were presented 5 different sentences and provided a rating after each presentation, which were were then aver-

 ©2000 American Medical Association. All rights reserved.

aged. The sentences were presented at an SPL of 52 dB, 62 dB, and 74 dB in a quiet background and then in the multitalker babble (S/B ratio, 10 dB).

Subjective Ratings

Two measures were used to elicit expressions of the quality of hearing aid performance from the patients. One was the Profile of Hearing Aid Performance/Profile of Hearing Aid Benefit (PHAP/PHAB), which quantifies 2 major aspects of hearing aid performance: speech communication in a variety of daily life situations and reactions to the loudness and quality of environmental sound.[33] For the PHAP/PHAB, 7 subscale scores were derived from the 66 items of the inventory that were completed by the patient in written format. The scales quantify problems in communication in favorable and unfavorable listening conditions as well as the aversiveness and distortion of a variety of sounds. The 7 subscales include communication with familiar talkers, ease of communication, reverberation, reduced cues, background noise, aversiveness of sound, and distortion of sound. At the end of each of the 3 trial periods, the patients completed the PHAP/PHAB inventory in the unaided and aided conditions using a 7-point scale that ranged from always to never. The PHAP scores quantify the scale scores in terms of aided performance, while PHAB scores quantify the scale scores in terms of benefit (ie, the difference between the aided scores and unaided scores). Hence, in the PHAP, scores for all subscales are reported in terms of percentage of time a problem is experienced and scores for the PHAB are reported in terms of the change in percentage of time a problem is experienced.

The second subjective assessment procedure was used at the final visit only. After having completed each of the 3 treatments, the patients provided, from memory, a rank-order rating of the 3.

Statistical Methods

A crossover design was chosen for this study instead of the more traditional randomized, parallel group design because it required fewer patients, eliminated between-patient variation, and it increased power for other objectives of the trial (eg, to determine which patient characteristics predict success with the different hearing aid circuits). In addition, some of the known disadvantages of the crossover design (eg, large dropout rate, instability of the patient's condition, and a large carryover effect) were not expected in this study. The 3 circuits were compared using aided scores and aided minus unaided scores (benefit scores) with a repeated measures analysis.

The sample of 360 patients provided at least 80% power to detect a small-to-medium effect size for the patients' rank-order rating among the 3 circuits. This sample size also provided greater than 95% power to detect a 7.2% difference in the NU-6 test, greater than 95% power to detect a 3.6% difference in the CST, greater than 90% power to detect a 20% difference in the Quality Rating Test, and 90% power to detect a 16.6% difference in the PHAB.

A mixed, repeated measures model was used to compare the 3 hearing aid circuits for the individual outcome variables. If the overall test was statistically significant, then pairwise comparisons were made between the groups using the Bonferroni procedure to adjust the α level for multiple tests. No adjustment was made for multiple outcomes. For this reason, P values close to .05 should be interpreted with caution. Sample sizes reported for specific tests and conditions departed somewhat from 360 because some patients did not complete the study, some were unable to perform the task, or, occasionally, the examiner was unable to follow the study's protocol.

RESULTS

Patient Sample

Four hundred forty-six patients were screened for inclusion in the trial and 360 (80.7%) were randomized. Of the patients who were not randomized, 15 were excluded on the basis of a single criterion, but most failed to meet 2 or more of the inclusion criteria. The main reasons included: air conduction thresholds exceeded 70 dB in either ear (20); a difference in pure tone averages between ears of more than 10 dB (17); mean air-bone gap exceeded 5 dB in either ear (13); or routine otoscopy did not reveal clear ear canals (13). In 7 instances, the audiologist did not feel that the patient was capable of performing the tasks required by the trial.

Of the 360 patients enrolled, 69.7% were military veterans. The mean age of the group was 67.2 years (range, 29-91 years). The racial/ethnic distribution approximated that of the US population: 78.6% were white; 12.2% black; 6.1% Hispanic; 1.9% Asian; and 1.1% Native American. Fifty-seven percent were men; women were mainly nonveteran patients who were authorized to be treated at VA medical centers for the purposes of this trial because the study grant funded the cost of the hearing aids and the time of the treating and evaluating audiologists. The most common self-reported causes of the patient's hearing loss were noise exposure and aging. About half (46.7%) had never used a hearing aid.

The number of patients from each center was nearly equal (range, 44-46). None of the groups representing the 6 randomized orders were statistically different in terms of age, age at onset of hearing loss, sex, race, previous hearing aid usage, and degree of hearing loss ($P \geq .11$ for all comparisons). Twenty-nine of the 360 patients did not complete the trial due to illness, relocation of residence, or other reasons (eg, withdrawal of patient consent, illness unrelated to hearing, death, sudden change in hearing). Three hundred thirty-seven patients completed the 90-day trial with the PC circuit, 338 with the CL, and 333 with the WDRC. The average reported hearing aid use time for the 3 circuits did not differ significantly and averaged about 9.8 (SD, 4) hours per day.

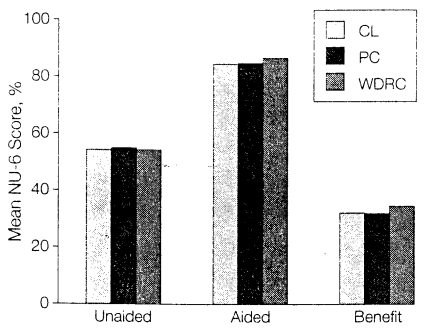
Speech Recognition Tests

FIGURE 3 provides a summary of the mean percentage correct results for the unaided and aided conditions for the

©2000 American Medical Association. All rights reserved.

NU-6 test for each of the 3 circuits together with the benefit scores (aided minus unaided). For statistical testing, percentage correct scores for the NU-6 test were arcsine transformed to stabilize the error variance.[34] Comparison of the unaided means with the aided means showed that each of the 3 circuits improved the mean word recognition score by a substantial amount (approximately 29% in absolute score differences; $P<.001$). The overall statistical test comparing the 3 circuits was significant ($P=.002$ for the aided scores and $P=.002$ for the benefit scores). Pairwise comparison tests showed that the WDRC circuit was superior to the other 2 circuits for the aided scores and superior to the PC circuit for the benefit scores.

Figure 3. Mean Percentage of Correctly Recognized Monosyllabic Words



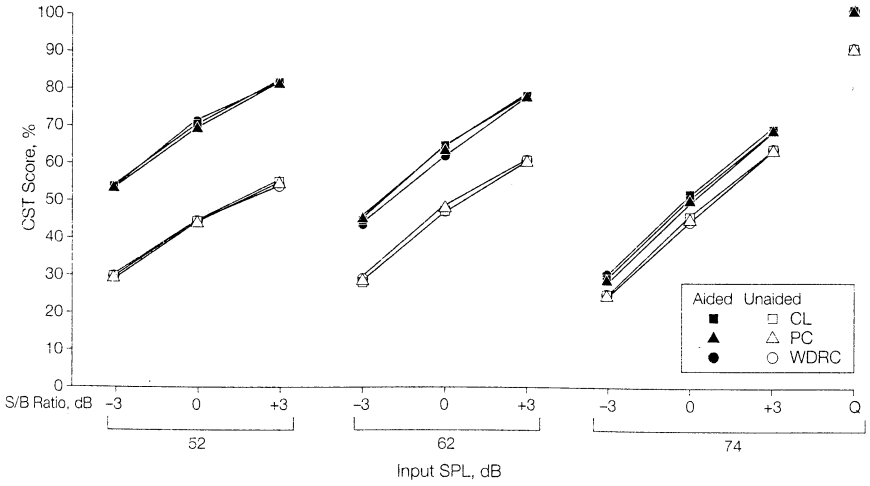
CL indicates compression limiter; PC, linear peak clipper; and WDRC, wide dynamic range compressor. The unaided values were obtained in the same test session as the aided values. Sample sizes varied slightly across conditions (320 to 324 for the aided conditions and 288 to 291 for the unaided conditions).

FIGURE 4 summarizes the findings for the aided and unaided CST results. Percentage correct scores were arcsine transformed to stabilize the error variance.[34] As expected, there were no differences among unaided means. However, significantly higher CST scores ($P<.001$) were achieved for all aided conditions relative to the unaided conditions. The overall statistical test comparing the 3 circuits for aided CST scores was significant for 1 condition (62/0; $P=.006$). Pairwise comparisons showed that the WDRC circuit was inferior to the CL and PC circuits.

The mean CST benefit scores (aided minus unaided) are shown in FIGURE 5. Comparison of the 3 circuits showed significant differences for the 62/0 ($P=.04$) and 74/0 conditions ($P=.02$). Pairwise comparison tests showed that for the 62/0 condition, the WDRC circuit was inferior to the CL circuit; and for the 74/0 condition, the WDRC circuit was superior to the CL circuit.

The data presented in Figures 4 and 5 also show that the 3 circuits provided similar amounts of improvement in test scores, but all showed successively less benefit as a function of signal level when background noise was present. A marked decrease in CST benefit scores from about 26% for the 52-dB conditions to approximately 6% for the 74-dB conditions was observed, suggesting that the hearing aids were less helpful at higher than at lower and moderate input levels. Furthermore, the Figures show that all 3 circuits provide measurable benefit in noisy conditions.

Figure 4. Connected Speech Recognition of Different Ratios of Signal to Ambient Noise



CL indicates compression limiter; PC, linear peak clipper; WDRC, wide dynamic range compressor; and SPL, sound pressure level. Mean unaided and aided scores on the Connected Speech Test (CST) shown for 10 test conditions for the 3 circuits. The abscissa labels the conditions by signal/babble levels (dB) for 9 conditions (52/-3 through 74/+3). S/B indicates signal-to-babble; Q, that the test was performed in quiet. Sample sizes varied across conditions (from 280 to 322 for unaided conditions and 320 to 336 for aided conditions).

Quality Rating Test

The Quality Rating Test was administered at 3 signal levels in quiet (designated as 52Q, 62Q, 74Q) and elicited ratings of loudness, noise interference, and overall liking. It was also administered at the same signal levels with an absolute S/B ratio of 10 dB (designated as 52N, 62N, and 74N), which means that the level of the speech was 10 dB greater than the level of the multitalker babble.

TABLE 1 shows no differences in the loudness ratings between the unaided means for each condition. Significant differences were observed, however, for the aided condition across the 3 circuits for both the quiet and background noise conditions for the lowest (52-dB SPL) and for the highest signal levels (74-dB SPL) ($P<.001$). The WDRC circuit was rated as being more comfortably loud (ie, a rating closer to 5) than the other 2 circuits for the 52-dB SPL conditions ($P=.003$) and 74-dB SPL conditions ($P=.003$). The CL circuit was more comfortably loud compared with the PC circuit for the 74-dB SPL condition.

A summary of data for the noise interference task is shown in TABLE 2. Analysis of the mean unaided data revealed no differences. For the aided data, the analysis also showed no significant differences among circuit types, except for the 62N condition ($P=.01$). Pairwise comparison revealed that the PC circuit scored higher (less noise interference) than the WDRC circuit.

©2000 American Medical Association. All rights reserved.

A summary of data for the overall liking task is shown in TABLE 3. There were no significant differences between the unaided means at each condition. The analyses of the data among circuit types for the aided condition showed significance for the 74Q condition ($P=.001$). Pairwise comparisons across circuits showed that the PC was less liked than both the CL and the WDRC.

Finally, for each circuit, significant improvement in overall liking was observed for soft and conversational speech levels ($P \leq .05$). For the loud conditions (74Q, 74N), however, negative average benefit ratings were observed ($P \leq .01$) except for the 74N condition for the WDRC ($P=.39$), suggesting that the aided experience was rated as being less liked than the unaided experience for loud sounds.

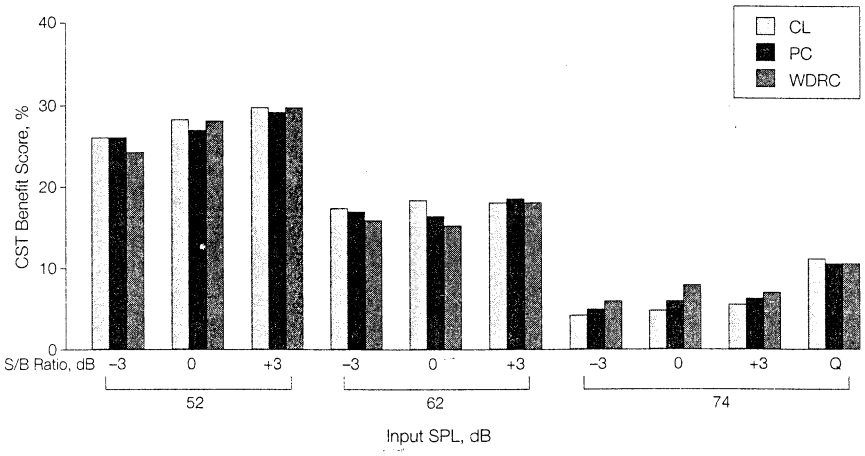
Subjective Assessment

No differences were observed among the unaided means for the PHAP. For the aided means, the analysis showed statistical significance ($P<.001$) for 2 of the 7 scales: distortion of sounds and aversiveness of environmental sounds. Pairwise comparisons showed that the scores for the PC were significantly different (ie, higher frequency of problems) than both the CL and WDRC circuits on the aversion and distortion scales ($P \leq .02$). The mean values for the PC circuit were 4% to 5% higher (ie, higher frequency of problems) for aversion and were 2% to 3% higher for distortion.

The PHAB scores also showed that each circuit significantly reduced the frequency of problems reported on 6 of the 7 scales ($P<.001$). For aversion, however, all circuits produced a significantly higher frequency of problems ($P<.001$) than in the unaided condition. In the analysis comparing circuit types, significant differences were observed for aversion ($P<.001$) and distortion ($P=.02$). Pairwise comparisons for aversion showed that the PC circuit was more aversive than both of the other circuits ($P=.003$) with the mean frequency of problems being 4% to 5% higher. The PHAB scores also showed that the PC score was significantly higher than the WDRC for distortion ($P=.02$) with the mean difference between PC and WDRC being 3%.

Finally, on completion of the study, the patients provided, from memory, a ranking of the 3 hearing aid circuits. The CL circuit received the highest percentage of first rankings (41.6%), followed by the WDRC (29.8%), and the PC (28.6%). In addition, the CL circuit was ranked third by the lowest percentage of patients (25.4% for the CL vs 36.2% for the PC and 38.4% for the WDRC).

Figure 5. Mean Benefit Scores on the CST of Different Ratios of Signal to Ambient Noise



Abbreviations are expanded in the legend of Figure 4. The abscissa labels the conditions as signal/babble level (dB) for 9 conditions (52/−3 through 74/+3). S/B indicates signal-to-babble; Q that the test was performed in quiet. Sample sizes across conditions varied from 277 to 321.

Table 1. Loudness Ratings Obtained in the Aided and Unaided Conditions*

		Aided			Unaided		
Condition, dB	Type of Circuit	No. of Subjects	Mean (SD)	*P* Value	No. of Subjects	Mean (SD)	*P* Value
52Q	CL	335	4.14 (0.94)	<.001†	318	2.77 (1.24)	.97
	PC	333	4.15 (1.00)		314	2.77 (1.31)	
	WDRC	330	4.43 (0.99)		312	2.78 (1.26)	
52N	CL	335	3.92 (1.06)	<.001†	318	2.46 (1.23)	.92
	PC	333	3.81 (1.08)		316	2.45 (1.24)	
	WDRC	328	4.21 (1.07)		312	2.43 (1.22)	
62Q	CL	335	5.33 (0.90)	.26	325	4.36 (1.21)	.48
	PC	333	5.34 (0.90)		325	4.31 (1.22)	
	WDRC	330	5.41 (1.00)		320	4.29 (1.16)	
62N	CL	334	5.31 (1.19)	.35	326	3.85 (1.27)	.73
	PC	333	5.20 (1.09)		325	3.86 (1.33)	
	WDRC	330	5.25 (1.31)		320	3.90 (1.37)	
74Q	CL	335	7.96 (1.55)	<.001‡	335	6.30 (1.63)	.84
	PC	333	8.31 (1.49)		332	6.28 (1.64)	
	WDRC	330	7.73 (1.59)		330	6.30 (1.67)	
74N	CL	335	7.57 (1.89)	<.001‡	335	5.88 (1.72)	.72
	PC	333	8.11 (1.69)		334	5.81 (1.74)	
	WDRC	330	7.26 (1.94)		330	5.87 (1.84)	

*CL indicates compression limiter; PC, linear peak clipper; and WDRC, wide dynamic range compressor.
†Pairwise comparison reveals PC and WDRC, and CL and WDRC are statistically significant.
‡Pairwise comparison reveals all 3 circuits are statistically significantly different.

©2000 American Medical Association. All rights reserved.

Statistical analysis using the Friedman test showed a significant overall difference among the rankings ($P=.002$). Subsequent analyses using the Wilcoxon test showed that, overall, the CL was preferred more frequently than the PC ($P=.001$) and the WDRC ($P=.001$) and that there were no significant differences between the rankings for the PC and the WDRC ($P=.86$).

Table 2. Noise Interference Ratings Obtained in the Aided and Unaided Conditions*

Condition, dB	Type of Circuit	Aided: No. of Subjects	Aided: Mean (SD)	Aided: *P* Value	Unaided: No. of Subjects	Unaided: Mean (SD)	Unaided: *P* Value
52Q	CL	335	9.74 (0.93)	.42	317	9.53 (1.33)	.16
	PC	332	9.80 (0.79)		314	9.57 (1.15)	
	WDRC	329	9.74 (0.86)		309	9.45 (1.46)	
52N	CL	336	7.10 (2.24)	.56	316	6.96 (2.43)	.99
	PC	332	7.18 (2.24)		315	6.93 (2.41)	
	WDRC	330	7.07 (2.19)		309	6.95 (2.47)	
62Q	CL	336	9.83 (0.66)	.86	323	9.66 (1.25)	.22
	PC	332	9.85 (0.53)		324	9.74 (0.95)	
	WDRC	328	9.83 (0.57)		318	9.66 (1.15)	
62N	CL	336	6.41 (2.28)	.01†	320	6.36 (2.28)	.62
	PC	333	6.55 (2.34)		322	6.37 (2.29)	
	WDRC	329	6.17 (2.38)		317	6.47 (2.23)	
74Q	CL	336	9.74 (0.73)	.92	332	9.67 (1.22)	.22
	PC	331	9.73 (0.85)		330	9.74 (1.07)	
	WDRC	329	9.72 (0.77)		327	9.66 (1.20)	
74N	CL	336	5.13 (2.47)	.69	334	5.48 (2.31)	.80
	PC	333	5.18 (2.48)		333	5.37 (2.36)	
	WDRC	329	5.25 (2.65)		329	5.43 (2.34)	

*CL indicates compression limiter; PC, linear peak clipper; and WDRC, wide dynamic range compressor.
†Comparison of PC circuit with WDRC circuit, which is statistically significantly different in pairwise comparison tests.

Table 3. Overall Liking Ratings Obtained in the Aided and Unaided Conditions*

Condition, dB	Type of Circuit	Aided: No. of Subjects	Aided: Mean (SD)	Aided: *P* Value	Unaided: No. of Subjects	Unaided: Mean (SD)	Unaided: *P* Value
52Q	CL	335	7.70 (2.36)	.14	316	5.63 (2.93)	.88
	PC	333	7.78 (2.39)		314	5.56 (2.80)	
	WDRC	330	7.96 (2.21)		308	5.57 (2.81)	
52N	CL	335	6.77 (2.38)	.64	316	4.86 (2.77)	.72
	PC	332	6.66 (2.45)		314	4.87 (2.74)	
	WDRC	330	6.76 (2.33)		309	4.75 (2.76)	
62Q	CL	336	8.44 (1.85)	.97	322	7.28 (2.53)	.28
	PC	331	8.43 (1.84)		323	7.31 (2.55)	
	WDRC	330	8.45 (1.78)		318	7.14 (2.51)	
62N	CL	335	6.85 (2.27)	.07	321	5.86 (2.55)	.90
	PC	330	6.97 (2.30)		323	5.82 (2.51)	
	WDRC	330	6.70 (2.30)		318	5.79 (2.60)	
74Q	CL	336	7.43 (2.60)	.001†	331	7.73 (2.31)	.08
	PC	332	6.91 (2.76)		330	7.94 (2.23)	
	WDRC	329	7.49 (2.54)		325	7.72 (2.24)	
74N	CL	336	5.80 (2.57)	.12	334	6.04 (2.39)	.70
	PC	331	5.59 (2.77)		333	6.05 (2.55)	
	WDRC	330	5.92 (2.64)		329	5.97 (2.47)	

*CL indicates compression limiter; PC, linear peak clipper; and WDRC, wide dynamic range compression.
†Comparison of CL circuit with PC circuit and comparison of PC circuit with WDRC circuit are statistically different in pairwise comparison tests.

COMMENT

Each of the 3 hearing aid circuits provides substantial benefit over unaided listening. Benefit was observed for measures of speech recognition and ratings of speech quality in a variety of noisy and quiet conditions as well as for subjective measures. Each circuit improved monosyllabic word recognition scores in a quiet background at conversational levels by an average of 29% (absolute score improvement). Speech recognition ability, as shown by the CST, in noise was improved by each circuit by amounts ranging from 10% to 30% (absolute score improvement), with greater improvement observed for speech at soft and conversational levels. Loudness rating data suggested that all 3 circuits amplified soft and conversationally loud speech to comfortable levels. The noise interference ratings showed that none of the circuits had a deleterious effect. For soft and conversational speech levels, each circuit improved the overall quality of the listening experience. For loud speech, the overall quality of listening was not significantly degraded. The results of 6 of the 7 subscales of the subjective measure of hearing aid benefit (PHAB) showed a significant reduction in the frequency of problems associated with communication in everyday environments.

Statistically significant differences (small in comparison with the benefits seen with each of the circuits) were noted among the circuits on several components of the outcome measures. The results of the loudness rating suggest that the WDRC circuit was more comfortably loud than the other 2 circuits for soft and for loud speech input conditions. Because of its operating characteristics, the WDRC was expected to produce a more comfortable listening experience for the soft and loud input levels. Differences in scores on the PHAP/PHAB for 2 subscales were statistically significant among circuits, with the PC rated as 4.5% more aversive than the other 2 circuits and producing an average of 3% more problems for distortion of sounds compared with the WDRC circuit. The preference rankings provided at the end of the trial favored the CL circuit. Because the dif-

©2000 American Medical Association. All rights reserved.

ferences between the hearing aid circuits were small in most cases, dispensing decisions should take into account cost vs benefit considerations for individual patients. In this regard, many programmable hearing aids (such as the one used in this trial) may be configured to function as a PC, CL, or WDRC and, as such, there are no cost differences between circuit options; however, for conventional, nonprogrammable devices, compression circuitry (either CL or WDRC) adds significantly to the single-unit price of the device.

Because concerted efforts were made to recruit patients into the study from both sexes and all racial groups, the study sample was a good representation of US adults who are candidates for hearing aids. We believe, therefore, that the study results are generalizable to the US population with sensorineural hearing loss. One limitation of the trial is that it did not measure other domains, such as affect and cognition, which are influenced by hearing loss.

Author Affiliations: Department of Veterans Affairs Medical Center, Washington, DC (Drs Larson, Beck, and Boysen); Hines VA Cooperative Studies Program Coordinating Center, Hines, Ill (Mr Williams and Dr Henderson); National Institute on Deafness and Other Communication Disorders, Bethesda, Md (Drs Luethke and Donahue); Greater Los Angeles VA Health Care System, Los Angeles, Calif (Dr Noffsinger); Department of Veterans Affairs Medical Center, Mountain Home, Tenn (Dr Wilson); University of Texas Health Sciences Center, San Antonio (Dr Dobie); Department of Veterans Affairs Medical Center, Iowa City, Iowa (Dr Haskell); Department of Veterans Affairs Medical Center, Nashville, Tenn (Dr Bratt); Department of Veterans Affairs Medical Center, Long Beach, Calif (Dr Shanks); Boys Town National Research Hospital, Omaha, Neb (Dr Stelmachowicz); University of Memphis, Memphis, Tenn (Dr Studebaker); Harborview/University of California, Los Angeles, Medical Center, Torrance (Dr Canalis); Department of Veterans Affairs Medical Center, Portland, Ore (Dr Fausti); and Department of Veterans Affairs Medical Center, Albuquerque, NM (Dr Rappaport).

NIDCD/VA Hearing Aid Clinical Trial Group: Vernon D. Larson, PhD,* chairman and principal investigator; Jillyn Roxberg, MA, clinical coordinator, and Rose Lee Jacob, program assistant, Veterans Affairs Medical Center, Mountain Home, Tenn; Joyce S. Kang, MS, clinical coordinator, and Megan M. Wilkinson, BA, program assistant, Veterans Affairs Medical Center, Washington, DC.

Participating Veterans Affairs Medical Centers, Participating Investigators and Audiologists: Albuquerque, NM: Bruce Z. Rappaport, PhD, Karen K. Stewart, MA, Sheree Smith, MA, and Jonni K. Dressendorfer, MS; Iowa City, Iowa: George B. Haskell, PhD, Kathy Vander Werff, MA, Adriana Rodriguez, MA, and Janette L. Rogers, MA; Long Beach, Calif: Janet E. Shanks, PhD, Charlene M. Mikami, MA, and Donna J. Gelnett, MA; Mountain Home, Tenn: Richard H. Wilson, PhD, Anne L. Oyler, MA, Kristen L. Bauer, MA, Colleen Noe, PhD, and Jennifer Crocket, MA; Nashville, Tenn: Gene W. Bratt, PhD,* Mia A. L. Rosenfeld, MS, Kjersten Branscome, MS, Mary Beth Trine, MS, and Barbara Peek, PhD; Portland, Ore: Stephen A. Fausti, PhD, James Henry, PhD, Wendy Helt, MA, Craig Larson, MA, and Alison Gilbert, MS; Washington, DC: Lucille B. Beck, PhD,* Gerald Schuchman, PhD, and Holly A. Sandvik, MS; and West Los Angeles, Calif: Douglas Noffsinger, PhD,* Donald D. Dirks, PhD, Diane Kenworthy, MA, Sheril A. Plunkett, MS, and Eleanor L. Wilson, MS.

Hines VA Cooperative Studies Program Coordinating Center, Hines, Ill: William G. Henderson, PhD,* director, Domenic Reda, MS, associate director, David Semlow, BS, assistant director, David W. Williams, MS,* study biostatistician, Mary Biondic, BA, programmer, and Barbara J. Christine, statistical assistant.

National Institute on Deafness and Other Communication Disorders, Bethesda, Md: James B. Snow, Jr, MD, retired director; Lynn E. Luethke, PhD,* health scientist administrator, and Amy Donahue, PhD,* chief, Hearing and Balance/Vestibular Sciences Branch.

Department of Veterans Affairs Headquarters, Washington, DC: John R. Feussner, MD, chief research and development officer, Joseph Gough, MA, program manager, and Ping C. Huang, PhD, staff assistant, Cooperative Studies Program; Lucille B. Beck, PhD,* director, and Allen E. Boysen, PhD,* retired director, Audiology and Speech Pathology Service.

Executive Committee: Rinaldo Canalis, MD, Torrance, Calif; Robert A. Dobie, MD, San Antonio, Tex; Patricia Stelmachowicz, PhD, Omaha, Neb; and Gerald A. Studebaker, PhD, Memphis, Tenn.

An asterisk indicates that a person is an Executive Committee member in the other sections above.

Data and Safety Monitoring Board: Noel Cohen, MD, chairman, New York, NY; Marilyn Demorest, PhD, Baltimore, Md; G. Pamela Burch-Sims, PhD, Nashville, Tenn; Wayne O. Olsen, PhD, Rochester, Minn; Richard A. Chole, MD, PhD, St Louis, Mo; Julianna Gulya, MD, Washington, DC; Debra L. Tucci, MD, Durham, NC; and Marie Diener-West, PhD, Baltimore, Md.

Funding/Support: Financial support was provided by the National Institute on Deafness and Other Communication Disorders and the Cooperative Studies Program of the Department of VA Research and Development Service.

REFERENCES

1. Collins JG. Prevalence of selected chronic conditions. *Vital Health Stat 10.* 1997;194:1-89.
2. Ries PW. Prevalence and characteristics of persons with hearing trouble: United States, 1990-1991. *Vital Health Stat 10.* 1994;188:1-75.
3. Kochkin S. MarkeTrak III. *Hear J.* 1992;45:47-54.
4. Goldstein DP. Hearing impairment, hearing aids and audiology. *ASHA.* 1984;26:24-35,38.
5. Leske MC. Prevalence estimates of communicative disorders in the US. *ASHA.* 1981;23:229-237.
6. Havlik RJ. Aging in the eighties, impaired senses for sound and light in persons age 65 and over. *Vital Health Stat 125.* 1986;125:1-8.
7. Mulrow CD, Aguilar C, Endicott JE, et al. Quality-of-life changes and hearing impairment: a randomized trial. *Ann Intern Med.* 1990;113:188-194.
8. Mulrow CD, Aguilar C, Endicott JE, et al. Association between hearing impairment and the quality of life of elderly individuals. *J Am Geriatr Soc.* 1990;38:45-50.
9. Uhlmann RF, Larson EB, Koepsell TD. Hearing impairment and cognitive decline in senile dementia of the Alzheimer's type. *J Am Geriatr Soc.* 1986;34:207-210.
10. Palmer CV, Adams SW, Bourgeois M, Durrant J, Rossi M. Reduction in caregiver-identified problem behaviors in patients with Alzheimer disease post-hearing-aid fitting. *J Speech Lang Hear Res.* 1999;42:312-328.
11. Russell JN, Hendershot GE, LeClere F, et al. Trends and differential use of assistive technology devices. *Vital Health Stat 292.* 1997;292:1-9.
12. Kochkin S. MarkeTrak III identifies key factors in determining consumer satisfaction. *Hear J.* 1992;45:39-44.
13. Kochkin S. Customer satisfaction and subjective benefit with high performance hearing aids. *Hear J.* 1996;49:16-26.
14. Cox RM, Alexander GC. The abbreviated Profile of Hearing Aid Benefit. *Ear Hear.* 1995;16:176-186.
15. Humes LE, Christensen LA, Bess FH, Hedley-Williams A. A comparison of the benefit provided by well-fit linear hearing aids and instruments with automatic reductions of low-frequency gain. *J Speech Lang Hear Res.* 1997;40:666-685.
16. Humes LE, Christensen L, Thomas T, Bess FH, Hedley-Williams A, Bentler R. A comparison of the aided performance and benefit provided by a linear and a two-channel wide dynamic range compression hearing aid. *J Speech Lang Hear Res.* 1999;42:65-79.
17. Dillon H, Birtles G, Lovegrove R. Measuring the outcomes of a national rehabilitation program. *J Am Acad Audiol.* 1999;10:67-79.
18. Hickson L. Compression amplification in hearing aids. *Am J Audiol.* 1994;3:51-65.
19. Dillon H. Compression? *Ear Hear.* 1996;17:4, 287-307.
20. Anderson BA, Hagen LT, Peterson TS, et al. Hearing instrument distortion. *Hear Rev.* 1996;69:18-26.
21. Hawkins DB, Naidoo SV. Comparison of sound quality and clarity with asymmetrical peak clipping and output limiting compression. *J Am Acad Audiol.* 1993;4:221-228.
22. Krebs D. Distortion and methods of output limiting. *Scand Audiol.* 1972;1:167-175.
23. Braida LD, Durlach NI, Lippmann RP, et al. Hearing aids—a review of past research on linear amplification, amplitude compression, and frequency lowering. *ASHA Monogr.* 1979;19:1-114.
24. Moore BCJ. Signal processing to compensate for reduced dynamic range. In: Beilin J, Jensen GR, eds. *Recent Developments in Hearing Instrument Technology, 15th Danavox Symposium, Kolding, Denmark, Mar 30-Apr 2, 1993.* Scanticon, Denmark: Danavox Jubilee Foundation; 1993:147-165.
25. Teder H. Compression in the time domain. *Am J Audiol.* 1993;2:41-46.
26. Plomp R. The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. *J Acoust Soc Am.* 1988;83:2322-2327.
27. Wilson RH. Development and use of auditory compact discs in auditory evaluation. *J Rehabil Res Dev.* 1993;30:342-351.
28. American National Standards Institute. *Specification of Hearing Aid Characteristics.* New York, NY: Acoustical Society of America; 1996. ANSI S3.22-1987.
29. Byrne D, Dillon H. The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear Hear.* 1986;7:257-265.
30. Hawkins DB, Walden BE, Montgomery A, et al. Description and validation of an LDL procedure designed to select SSPL90. *Ear Hear.* 1987;8:162-169.
31. Cox RM, Alexander GC, Gilmore C. Development of the Connected Speech Test (CST). *Ear Hear.* 1987;8(5 suppl):119S-126S.
32. Cox RM, Alexander GC, Gilmore C, Pusakulich KM. Use of the Connected Speech Test (CST) with hearing-impaired listeners. *Ear Hear.* 1988;9:198-207.
33. Cox RM, Gilmore C. Development of the Profile of Hearing Aid Performance (PHAP). *J Speech Lang Hear Res.* 1990;33:343-357.
34. Studebaker GA. A "rationalized" arcsine transform. *J Speech Lang Hear Res.* 1985;28:455-462.

©2000 American Medical Association. All rights reserved.