Dictated versus database-generated discharge summaries: a randomized clinical trial

Evidence

Études

Carl van Walraven,*† MD, MSc; Andreas Laupacis,*† MD, MSc; Ratika Seth,* MD; George Wells,*† PhD

From *the Department of Medicine and †the Clinical Epidemiology Unit at the Loeb Research Institute, University of Ottawa, Ottawa, Ont.

Dr. van Walraven was an R. Samuel McLaughlin Foundation research fellow at the Institute for Clinical Evaluative Sciences, Toronto, Ont.

This article has been peer reviewed.

CMAJ 1999;160:319-26

? See related article page 345

Abstract

Background: Hospital discharge summaries communicate information necessary for continuing patient care. They are most commonly generated by voice dictation and are often of poor quality. The objective of this study was to compare discharge summaries created by voice dictation with those generated from a clinical database.

Methods: A randomized clinical trial was performed in which discharge summaries for patients discharged from a general internal medicine service at a tertiary care teaching hospital in Ottawa were created by voice dictation (151 patients) or from a database (142 patients). Patients had been admitted between September 1996 and June 1997. The trial was preceded by a baseline cohort study in which all summaries were created by dictation. For the database group, information on forms completed by housestaff was entered into a database and collated into a discharge summary. For the dictation group, housestaff dictated narrative letters. The proportion of patients for whom a summary was generated within 4 weeks of discharge was recorded. Physicians receiving the summary rated its quality, completeness, organization and timeliness on a 100-mm visual analogue scale. Housestaff preference was also determined.

Results: Patients in the database group and the dictation group were similar. A summary was much more likely to be generated within 4 weeks of discharge for patients in the database group than for those in the dictation group (113 [79.6%] v. 86 [57.0%]; $p < 0.001$). Summary quality was similar (mean rating 72.7 [standard deviation (SD) 19.3] v. 74.9 [SD 16.6]), as were assessments of completeness (73.4 [SD 19.8] v. 78.2 [SD 14.9]), organization (77.4 [SD 16.3] v. 79.3 [SD 17.2]) and timeliness (70.3 [SD 21.9] v. 66.2 [SD 25.6]). Many information items of interest were more likely to be included in the database-generated summaries. The database system created summaries faster and was preferred by housestaff. Dictated summaries in the baseline and randomized studies were similar, which indicated that the control group was not substantially different from the baseline cohort.

Interpretation: The database system significantly increased the likelihood that a discharge summary was created. Housestaff preferred the database system for summary generation. Physicians thought that the quality of summaries generated by the 2 methods was similar. The use of computer databases to create hospital discharge summaries is promising and merits further study and refinement.

Résumé

Contexte : Les résumés de congé d'hôpital fournissent de l'information nécessaire au soin continu des patients. Ils sont le plus souvent dictés et, souvent, de qualité médiocre. Cette étude visait à comparer des résumés de congé dictés à des résumés produits à partir d'une base de données cliniques.

Méthodes : On a réalisé une étude clinique randomisée au cours de laquelle les résumés de congés de patients libérés d'un service de médecine interne générale d'un hôpital d'enseignement de soins tertiaires à Ottawa ont été dictés (151 pa-

tients) ou produits à partir d'une base de données (142 patients). Les patients avaient été admis entre septembre 1996 et juin 1997. L'étude a été précédée d'une étude de cohortes de base au cours de laquelle tous les résumés ont été dictés. Dans le cas des résumés produits à partir de la base de données, on a entré dans une base de données les renseignements contenus sur les formules remplies par le personnel, et on les a réunis en un résumé de congé. Dans le cas des résumés dictés, des membres du personnel ont dicté des lettres narratives. On a pris note de la proportion des patients sur lesquels un résumé a été produit dans les quatre semaines suivant le congé. Les médecins qui ont reçu le résumé en ont évalué la qualité, l'intégralité, l'organisation et la pertinence sur une échelle analogique et visuelle de 100 mm. On a aussi déterminé la préférence du personnel de l'hôpital.

Résultats : Les patients dont le résumé provenait de la base de données et ceux dont le résumé avait été dicté étaient semblables. Les patients du premier groupe étaient beaucoup plus susceptibles que ceux du deuxième de faire l'objet d'un résumé produit dans les quatre semaines suivant leur congé (113 [79,6 %] c. 86 [57,0 %]; $p < 0{,}001$). La qualité des résumés était semblable (évaluation moyenne de 72,7 [écart type (ET) de 19,3] c. 74,9 [ET de 16,6], et les évaluations relatives à l'intégralité (73,4 [ET de 19,8] c. 78,2 [ET de 14,9]), à l'organisation [77,4 [ET de 16,3] c. 79,3 [ET de 17,2]) et à la pertinence (70,3 [ET de 21,9] c. 66,2 [ET de 25,6] se ressemblaient aussi. Les résumés produits à partir de bases de données étaient beaucoup plus susceptibles de comporter de nombreux éléments d'information d'intérêt. Le système de base de données a créé des résumés plus rapidement et avait la préférence du personnel de l'hôpital. Les résumés dictés au cours de l'étude de référence et de l'étude randomisée étaient semblables, ce qui indique que le groupe témoin n'était pas très différent de la cohorte de référence.

Interprétation : Le système de base de données a augmenté considérablement la probabilité qu'un résumé de congé soit créé et c'est le système que préférait le personnel de l'hôpital. Les médecins étaient d'avis que les résumés produits par les deux méthodes étaient de qualité semblable. L'utilisation des bases de données informatisées pour créer des résumés de congés d'hôpital se révèle prometteuse et mérite d'être étudiée plus à fond et perfectionnée.

Hospital discharge summaries are commonly used to communicate information between hospital-based and community physicians.[1] Previous studies have shown deficiencies in discharge summary content,[2–4] accuracy[5] and timeliness.[1,4,6,7]

Interventions to improve discharge summaries that have been tested in clinical studies include physician education,[8] handwritten interim reports[9–12] and standardization of the summary's format.[13–18] To make voice dictation unnecessary, clinical databases have been used to generate discharge summaries.[19–33] Studies have suggested that with a database method, the likelihood of discharge summary generation is greater,[28,34,35] summary accuracy is greater[28] and summaries are created more quickly.[34,35] Database summaries are also preferred by community physicians.[36–38] However, these studies may be unreliable, for several reasons. In few of the studies did the patient's own physician measure summary quality. In most of the studies a limited number of outcomes were reported, and in many the investigators did not compare dictated and database-generated summaries concurrently. The method of summary generation was not randomly assigned in any of the studies, and none of the investigators determined whether summaries in the control group changed after the database method was introduced.

Since it is unclear whether database-generated discharge summaries are better than dictated summaries, we conducted a randomized clinical trial to compare hospital discharge summaries created from a clinical database with those generated by voice dictation. To determine whether the dictated summaries created during the trial differed from those generated before the trial began, we first performed a baseline study in which all summaries were generated by dictation.

Methods

In this report physicians-in-training who treated inpatients are called "housestaff" and their supervisors are called "staffpeople." Physicians caring for patients after discharge from hospital are called "community physicians."

The study took place between September 1996 and June 1997 on the general internal medicine service at the Ottawa Civic Hospital, a 700-bed tertiary care teaching hospital. The hospital's institutional review board approved the study. The 80-bed service consisted of 4 clinical teams composed of a staffperson certified by the Royal College of Physicians and Surgeons of Canada and housestaff (a second- or third-year internal medicine resident and 1 or 2 interns and medical students). Throughout the study all housestaff received individual information sessions during which the study was described, methods to optimize the quality of dictated summaries were reviewed, and, during the randomized trial, the discharge summary database was explained.

All patients admitted to the general internal medicine service during the study period were eligible for inclusion. Patients were excluded if they had been transferred to or from another service, died during their hospital stay or remained in hospital past the dates specified by the study protocol. During the randomized trial, patients were also excluded if they had been admitted without a standardized history and physical admission form.

To create a dictated discharge summary, housestaff recited a letter into the hospital dictation system. The dictating housestaff decided what information was included and how it was organized. Dictations were transcribed in the medical records department, and copies of the summary were sent to all community physicians cited by the housestaff. For most physicians within the city, summaries were received within 3 working days of dictation.

Database fields were chosen because they led to significantly better summary quality.[39] Fields were grouped into preadmission, hospital and discharge information and corresponded to fields on 3 separate forms (the admission history and physical examination form, the hospital summary form and the database "Dear Doctor" letter, respectively). Housestaff completed these forms during the patient's hospital stay. The day after discharge, information from the 3 completed forms was entered verbatim into the database by the principal investigator (C.v.W.). If a form was blank, a sticker instructing housestaff to complete it was placed on the form. After data entry, a word processor macro command was used to collate the database information into a database discharge summary.

The unit of randomization was the hospital admission. Balanced randomization stratified by the physician team was used. Random number tables were used to generate block sizes and randomization schedules,[40,41] to which the investigators were blinded.

If a patient was assigned to the database group, the hospital summary form and database "Dear Doctor" letter were placed in his or her chart. Housestaff were informed that a patient had been assigned to the dictation group by means of a form similar to the hospital summary form. Instead of having database fields, this form reminded housestaff to dictate the summary when they discharged the patient, to send a copy of the summary to all involved physicians, and to include only information they thought was necessary for continuing patient care. Instead of the database "Dear Doctor" letter, the routinely used interim discharge letter was placed in the chart.

The primary outcomes included the proportion of admissions for which a discharge summary was created by 4 weeks after discharge and overall summary quality. The summaries were rated by community physicians on a 100-mm visual analogue scale ranging from 0 (worst) to 100 (best). A high-quality summary was defined as one that "efficiently communicates information necessary for continued patient care." The study's secondary outcomes, including summary completeness ("All necessary information was included"), organization and timeliness ("time from patient discharge to summary receipt") were also rated by the community physicians. The assessment form, along with an addressed, stamped envelope, was sent to the community physicians with the discharge summary. If the forms were not returned within 2 months of summary transcription, a reminder questionnaire and summary were sent. The assessment form was found to discriminate between summaries of seemingly different quality and to detect changes in summary quality in 2 pilot surveys involving 315 physicians (data available from the authors on request).

Additional measures of these outcomes were recorded. Chart reviews identified all consultations, procedures, medical therapies, complications and specific laboratory and radiology tests (Appendix 1). To measure summary completeness, we determined whether information found at chart review was cited in the summary. In addition, explicit criteria (available from the authors on request) were used to quantify other information cited in the summary. The summaries were reviewed by 1 of 2 investigators (C.v.W. and R.S.). Interrater reliability, based on 45 summaries, was high (overall κ 0.70, 95% confidence interval 0.65 to 0.76). Summary organization was measured as the proportion of content items reported with a heading or cited in the first sentence of a paragraph. Finally, timeliness was recorded as the number of days from patient discharge to summary generation.

Housestaff opinions were measured by means of a mail survey at the end of the study. Housestaff rated the time required and the burden of summary generation from 0 (worst) to 100 (best). They indicated their preference for either method of summary generation on a 100-mm visual analogue scale. Ratings below 50 indicated a preference for the dictation system, and ratings above 50 indicated a preference for the database system.

With the exception of summary assessment forms and housestaff surveys, data collection was complete for each patient.

Sample size calculation was based on summary quality. A planned interim analysis of the baseline study provided the mean and variance of dictated summary quality and variables needed for sample size adjustment (including the probability that a summary would not be generated, that a completed summary assessment form would not be returned or that the patient would die in hospital). The median minimal important difference from a pilot survey of 77 family physicians was a relative improvement of 19% in summary quality. Since the interim summary quality rating on the visual analogue scale was 74.3, the trial needed enough power to detect a mean quality rating of 88.4 for the database-generated summaries. This was

decreased to 85 to ensure that an important difference was not missed. With an α-error of 0.05 (2-sided) and a β-error of 0.15 (1-sided), the unadjusted total sample necessary to detect this difference was 94 summary quality assessments. After adjustments, we calculated that 372 patients needed to be included in the randomized trial to detect this difference.

Summaries created more than 4 weeks after discharge (16.4% of the summaries) were excluded from analysis since these summaries are usually different from those generated closer to patient discharge in purpose, content and length (personal observation). Study results did not change when these summaries were included. When more than one assessment form was received for a single summary, the mean rating for quality, completeness, organization and timeliness was used.

We compared continuous measures using Student's *t*-test. Student's paired *t*-test was used to compare housestaff opinions. Categorical measures were compared by means of the χ^2 test. Kaplan–Meier plots of time to summary creation were compared with the log-rank test. A 2-tailed *p* value less than 5% was considered significant for all analyses, which were performed with SPSS for Windows (version 7.0, SPSS Inc., Chicago).

Results

Randomized trial profile

Patient flow through the randomized trial is shown in Fig. 1. The 3 patients in the database group not eligible for summary generation for other reasons were excluded because the hospital summary form was accidentally removed by administrative staff during chart "thinning" (1 case), a housestaff member who had just started in the general internal medicine service dictated a summary for a patient in the database group (before the information session) (1 case), and a housestaff member incorrectly assumed that a "stat" database summary could not be generated (1 case). Therefore, the dictation and database groups had data for 151 and 142 patients respectively available for analysis.

The 2 groups were similar except that patients in the database group were more likely than those in the dictation group to be in a monitored bed (22.5% v. 11.9%; $p = 0.02$) (Table 1). The groups were also similar when only patients for whom a discharge summary was actually generated were compared.

A summary was much more likely to be generated within 4 weeks of discharge for patients in the database group (113 [79.6%] v. 86 [57.0%]) ($p < 0.001$). This large difference remained when the 4-week "deadline" imposed by the study's protocol was removed and when data for patients whose length of stay was less than 2 days were excluded. The groups did not differ with respect to the summary author's level of training. The proportion of summaries in the 2 groups assessed by at least one community physician was similar.

Physician assessments

During the randomized trial, assessments were available for 210 (69.5%) of 302 summaries (note that one summary could be sent to more than one physician). The summaries in the dictation and database groups were similar for all outcomes, including quality, even when adjusted for monitored bed status (Fig. 2). However, ratings differed significantly when stratified by the community physician's specialty. Family physicians gave higher timeliness ratings to database-generated summaries than to dictated summaries (mean rating 72.2 [standard deviation (SD) 22.7] v. 62.6 [SD 28.2]; $p = 0.04$). Consultants gave lower ratings to database summaries than to dictated summaries for quality (mean rating 64.6 [SD 24.7] v. 76.2 [SD 13.6]; $p = 0.02$) and completeness (mean rating 68.2 [SD 21.7] v. 79.5 [SD 13.5]; $p = 0.01$).

Other assessments

The completeness of the database and dictated summaries is shown in Table 2. Fifteen items were cited with significantly different frequency in the 2 groups. Ten of these items (including discharge diagnosis, discharge medications and planned follow-up) were more commonly cited in the database summaries than in the dictated summaries, whereas dictated summaries were more likely to list the social history, admission diagnosis, hospital consultations and functional status at discharge.

Database-generated summaries were shorter than dictated summaries (mean 57.3 [SD 17.1] lines v. 64.8 [SD 26.6] lines; $p = 0.03$) and placed 8 of 14 content items (including chief complaint, admission medications, physical examination, treatment, complications, pending laboratory results and recommendations) under their own headings more frequently. Considering only patients for whom a summary was created within 4 weeks of discharge, database-generated summaries were produced more quickly: 94.7% of the summaries in this group were generated within 1 week after discharge, compared with 80.2% in the dictation group (log-rank statistic 72.56, $p < 0.001$).

The mail survey was completed by 37 (86%) of 43 housestaff. Housestaff thought the database method was significantly faster than the dictation method (mean rating 65.3 [SD 19.8] v. 46.3 [SD 22.4]; $p = 0.007$) and less burdensome (mean rating 65.2 [SD 20.0] v. 43.0 [SD 22.4]; $p = 0.002$). The database method was significantly preferred (mean preference rating 70.4 [SD 23.4]; $p < 0.001$).

Comparison of dictated summaries from baseline and randomized studies

A total of 960 patients were discharged from the general internal medicine service during the baseline study. Fifty-

seven patients were ineligible for the study because they had been transferred from other services (28 patients), had been discharged from the emergency department (13 patients) or had not been treated by housestaff (16 patients). Of the 903 remaining patients, 121 died in hospital, which left 782 admissions eligible for generation of a dictated summary.

Patients in the baseline study and the randomized study were similar except that those in the latter study had a shorter stay, were less likely to be in a monitored bed and were less likely to have circulatory disease. These differences persisted when only patients for whom a summary was generated were compared. The probability that a dictated summary would be generated by 4 weeks after discharge was similar for the baseline and randomized groups (404/782 [51.7%] v. 86/151 [57.0%]). The quality, completeness, organization and timeliness of the dictated summaries were similar between the 2 studies (Fig. 2). Summaries from the randomized trial were more likely to cite hospital procedures and medical therapy, whereas those from the baseline study were more likely to cite active past medical history, social history and functional status at discharge. The proportion of summaries generated by 1 week

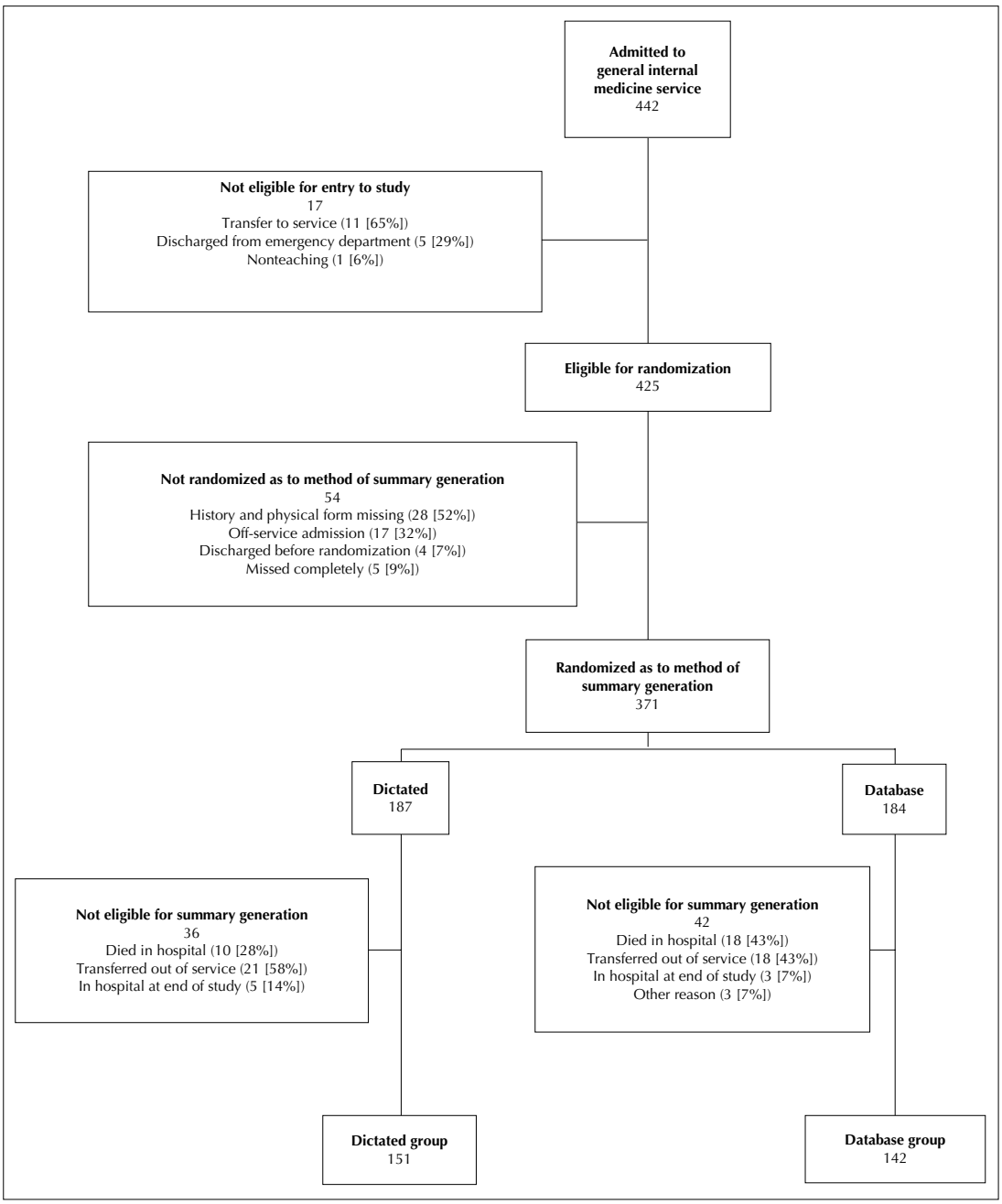

Fig. 1: Patient flow through randomized study comparing hospital discharge summaries created by voice dictation with those generated from a clinical database.

after discharge was 62.1% in the baseline study and 80.2% in the randomized trial (log-rank statistic 5.33, $p < 0.05$).

Interpretation

The likelihood that a summary would be generated was significantly greater with the database system than with the dictation system. Community physicians considered the database and dictated summaries to be similar. The database summaries contained more information, yet were shorter and were created more quickly. Housestaff preferred the database method of summary generation. The dictated summaries generated during the randomized trial were not extensively different from those generated during the baseline study.

The better summary generation with the database system has the potential to increase communication between hospital-based and community physicians. This could lead to better continuity of care and less duplication of health care services. Summaries were more likely to be generated with the database because housestaff found it quicker and easier — 2 of the 3 forms necessary for the database were completed during regular patient care. Therefore, with minimal extra work, the database summary made dictation unnecessary.

Physicians found that the database and dictated summaries were of similar quality. Given the 95% confidence limits around the difference between the groups, it is unlikely that their quality differed by more than 7.5 mm on a 100-mm scale.[42,43] In our prestudy survey only 18 (25%) of 72 respondents chose a minimal important difference of less than 7.5 mm. Therefore, if physicians in this nonrandom survey are representative of physicians in our community, 75% would consider database-generated and dictated summaries to be of equal quality.[44]

Specialists and family physicians had different views when database and dictated summaries were compared. Specialists gave database summaries lower ratings for completeness and overall quality. Since specialists often dictate narrative consultation notes, they may prefer the dictated summary because its format is more familiar to them.

Database-generated summaries were more likely to con-

Table 1: Preadmission and in-hospital characteristics of patients for whom discharge summaries were either created by voice dictation or generated from a database

Characteristic	Group: Dictation *n = 151*	Group: Database *n = 142*	*p* value*
Mean age (and SD), yr	62.9 (18.5)	65.7 (17.5)	0.18
% female	46.4	52.8	
Mean no. of preadmission diagnoses (and SD)	2.6 (1.7)	2.8 (2.1)	
Preadmission diagnosis, no. (and %) of patients			
Hypertension	33 (21.9)	42 (29.6)	0.13
Coronary artery disease	27 (17.9)	32 (22.5)	
Congestive heart failure	21 (13.9)	16 (11.3)	
Asthma or chronic lung disease	26 (17.2)	25 (17.6)	
Diabetes mellitus with or without complications	33 (21.9)	29 (20.4)	
Cancer	17 (11.3)	9 (6.3)	0.14
Mean length of stay (and SD), d	5.5 (5.4)	5.6 (4.7)	
Mean no. of new therapies (and SD)	2.40 (1.9)	2.30 (1.9)	
In-hospital characteristics, no. (and %) of patients			
Admitted to monitored bed	18 (11.9)	32 (22.5)	0.02
With at least 1 extreme laboratory result†	82 (54.3)	81 (57.0)	
With at least 1 diagnostic test†	132 (87.4)	132 (93.0)	0.11
With at least 1 consultation	86 (57.0)	78 (54.9)	0.12
With at least 1 complication	25 (16.6)	27 (19.0)	
With at least 1 procedure	59 (39.1)	46 (32.4)	
Mean no. of discharge medications (and SD)	3.82 (2.7)	3.90 (2.9)	
Primary discharge diagnosis, no. (and %) of patients			0.18
Respiratory system disorders	31 (20.5)	24 (17.0)	
Circulatory system disorders	16 (10.6)	30 (21.3)	
Digestive system disorders	30 (19.9)	31 (22.0)	
Injuries/poisonings/ undefined	12 (7.9)	12 (8.5)	
Neoplastic/blood disorders	20 (13.2)	10 (7.1)	
Endocrine/nutritional disorders	11 (7.3)	8 (5.7)	

Note: SD = standard deviation
*Provided only if less than 0.2.
†See Appendix 1 for definition.

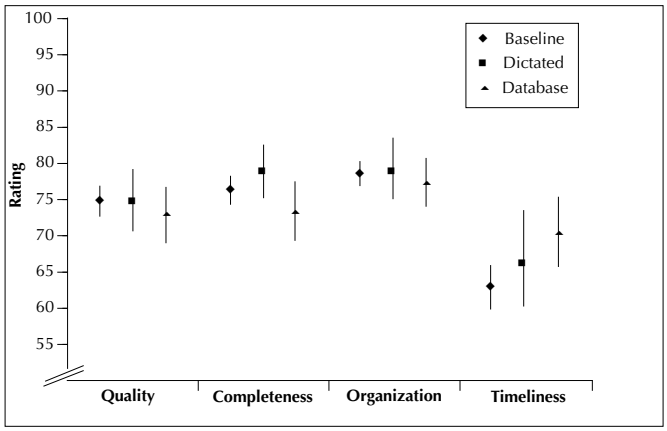

Fig. 2: Community physicians' assessments of discharge summaries for patients in the baseline study (all dictated [DIC-B]), those with dictated summaries in the randomized trial (DIC) and those with database-generated summaries in the randomized trial (DB). Summaries were rated on a visual analogue scale ranging from 0 (worst) to 100 (best). Error bars represent 95% confidence intervals.

tain many content items. We believe this is because the database forms completed by housestaff prompted them for this information, making data omission less likely. Many of the content items that were more commonly cited in the database summaries, including discharge diagnosis, discharge medications and patient follow-up care, have been identified in several surveys as important for discharge summary quality.[36,37,39,45] However, we are unsure what effect this change in summary content would have on patient care.

Other direct comparisons of database-generated and dictated summaries have shown results similar to ours. Lissauer and colleagues[28] found that their database significantly increased summary generation and improved reporting of diagnostic and content items. The family physicians in the survey by Brazy and associates[38] found that database summaries were faster, easier to abstract data from and more informative. Smith and Holzman[35] reported that database summaries were generated more quickly. Although other studies have shown that database summary systems have advantages over summary dictation, variability between these studies in terms of patients, databases and quality makes further generalizations hazardous.

Table 2: Discharge summary completeness*

Item	Group; no. (and %) of summaries with item: Dictated *n = 86*	Database *n = 113*	*p* value†
Chief complaint	84/86 (97.7)	113/113 (100)	0.19
History of presenting illness	84/86 (97.7)	109/113 (96.5)	
Active past medical history	72/83 (86.7)	106/106 (100)	0.001
Social history	32/86 (37.2)	7/113 (6.2)	0.001
Preadmission medications	57/86 (66.3)	113/113 (100)	0.001
Results of physical examination at admission	75/86 (87.2)	112/113 (99.1)	0.001
Admission diagnosis	61/86 (70.9)	1/113 (0.9)	0.001
Consultations‡	37/79 (46.8)	19/100 (19.0)	0.001
Procedures‡	37/48 (77.1)	33/46 (71.7)	
New medical therapy‡	123/239 (51.5)	128/262 (48.8)	
Complications‡	8/19 (42.1)	16/27 (59.2)	
Extreme result of blood testing‡	41/136 (30.1)	63/130 (48.5)	0.002
Results of diagnostic blood tests‡	19/114 (16.7)	39/130 (30.0)	0.01
Results of radiology tests‡	90/230 (39.1)	112/238 (47.0)	0.08
Discharge diagnosis	56/86 (65.1)	113/113 (100)	0.001
Discharge medications	80/86 (93.0)	113/113 (100)	0.006
Medical follow-up	76/80 (95.0)	104/105 (99.0)	0.57
Community services	18/86 (20.9)	46/113 (40.7)	0.003
Discharge functional status	9/86 (10.5)	1/113 (0.9)	0.003
Pending laboratory tests	8/86 (9.3)	46/113 (40.7)	0.001
Recommendations	33/86 (38.4)	61/113 (54.0)	0.03

*The denominators vary because not all items applied to all summaries.
†Provided only if less than 0.2.
‡The occurrence of these items was determined by a review of medical and laboratory records.

Several factors may explain why dictated discharge summaries were slightly different during the randomized trial from those in the baseline study. First, all patients in the randomized trial had forms placed in their chart reminding housestaff that the summaries would be assessed.[46] Second, forms for patients in the dictation group contained reminders that might have improved summary dictation. Finally, summaries generated by dictation during the randomized trial may have been contaminated by the database method.

Our study has several strengths compared with previous assessments of discharge summary interventions. First, patients were randomly assigned to 1 of the 2 methods of summary generation. Second, the inclusion criteria were broad and objective, with 371 (87.3%) of 425 eligible patients being included. Third, the dictated summary was studied extensively before the introduction of the database system and the randomized trial. We can therefore conclude that the equivalent summary quality in the randomized trial did not result from improvements in the dictated summary. Fourth, one of the primary outcomes was the community physicians' assessment of quality. Thus, the views of physicians who used the summary for continuing patient care were measured. Finally, the outcomes chosen for the study were comprehensive.

Our study had some limitations. First, housestaff were not blinded to the intervention, which made cointervention and contamination between groups possible. However, since the dictated summaries in the baseline study and the randomized trial were similar, major contamination was unlikely. Second, a summary was not generated for all patients, 9.4% of summaries were not sent to eligible physicians, and not every physician returned the summary assessment form. Therefore, intention-to-treat analysis was not possible for the summary assessment. However, all patients in the randomized trial were considered for summary generation, and all summaries were assessed by using other study outcomes. Third, since the study included only general internal medicine patients and housestaff at a single teaching hospital, the results may not be generalizable to other services or nonteaching centres. Finally, during the randomized study, completed database forms were entered into the database by the primary investigator. In a nonstudy setting, data would have to be entered by other workers, such as health records analysts. Since the investigator directly transcribed data from the completed forms, we believe that health records analysts could produce similar database summaries.

Competing interests: None declared.

References

1. Long A, Atkins JB. Communications between general practitioners and consultants. *BMJ* 1974;4:456-9.
2. Tulloch AJ, Fowler GH, McMullan JJ, Spence JM. Hospital discharge reports: content and design. *BMJ* 1975;4:443-6.
3. Mageean RJ. Study of "discharge communications" from hospital. *BMJ (Clin Res Ed)* 1986;293:1283-4.
4. van Walraven C, Weinberg AL. Quality assessment of a discharge summary system. *CMAJ* 1995;152(9):1437-42.
5. Macaulay EM, Cooper GG, Engeset J, Naylor AR. Prospective audit of discharge summary errors. *Br J Surg* 1996;83:788-90.
6. Geitung JT, Kolstrup N, Fugelli P. Written information from hospital to primary physician about discharged patients. *Tidsskr Nor Laegeforen* 1990; 110(24):3132-5.
7. Haikio JP, Linden K, Kvist M. Outcomes of referrals from general practice. *Scand J Prim Health Care* 1995;13(4):287-93.
8. Flyer B, Rubenstein LZ, Robbins AS, Wieland GD, Henry D, Cugalj N. An intervention to improve the hospital discharge summary. *J Med Educ* 1988;63(5):407-9.
9. Penney TM. Delayed communication between hospitals and general practitioners: Where does the problem lie? *BMJ* 1988;297:28-9.
10. Kendrick AR, Hindmarsh DJ. Which type of hospital discharge report reaches general practitioners most quickly? *BMJ* 1989;298:362-3.
11. Colledge NR, Smith RG, Lewis SJ. The delivery of interim discharge summaries to general practitioners by the elderly. *Health Bull* 1992;50(3):219-22.
12. Sandler DA, Mitchell JR. Interim discharge summaries: How are they best delivered to general practitioners? *BMJ (Clin Res Ed)* 1994;1987:1523-5.
13. Oetgen WJ, Steinweg DL, Kimball DB. The hospital discharge summary: a new format. *Milit Med* 1982;147(12):1054-5.
14. Howard DJ. Structured discharge letter in a department of geriatric medicine. *Health Trends* 1986;18(1):12-4.
15. King MH, Barber SG. Towards better discharge summaries: brevity and structure. *West Engl Med J* 1991;106(2):40-1.
16. Lloyd BW, Barnett P. Use of problem lists in letters between hospital doctors and general practitioners. *BMJ* 1993;306:247.
17. Rawal J, Barnett P, Lloyd BW. Use of structured letters to improve communication between hospital doctors and general practitioners. *BMJ* 1993;307:1044.
18. van Walraven C, Duke SM, Weinberg AL, Wells PS. Family physicians prefer standardized discharge summaries: a survey. *Can Fam Physician* 1998;44: 62-9.
19. Roger FH, Joos M, Haxhe JJ. An interhospital problem-oriented and automated discharge summary. *Methods Inf Med* 1976;15:141-4.
20. Jelovsek F, Smith R, Blackmon L, Hammond W. Computerized nursery discharge summary. *Methods Inf Med* 1977;16:199-204.
21. Janik DS, Swarner OW, Henriksen KM, Wyman ML. A computerized single entry system for recording and reporting data on high-risk newborn infants. *J Pediatr* 1978;93:519-23.
22. Stern PH, Rubin EH. Computerized discharge summaries. *Arch Phys Med Rehabil* 1979;60(1):25-9.
23. Maresh M, Beard RW, Combe D, Dawson AM, Gillmer MD, Smith G, et al. Selection of an obstetric data base for a microcomputer and its use for on-line production of birth notification forms, discharge summaries and perinatal audit. *Br J Obstet Gynaecol* 1983;90:227-31.
24. Finer NN, Fraser AJ. Neonatal data base and automated discharge summary using a personal computer and proprietary software. *Pediatrics* 1985;76:269-73.
25. van Domburg RT, Simoons ML. Computer-generated discharge letters in the coronary care unit. *Med Inf (Lond)* 1985;10(3):259-66.
26. Castleden WM, Lawrence-Brown MM, Lam HS, McLoughlin BS, Thompson DS, Lopez J. The Hollywood surgical-audit programme: a computer-based discharge and data-collection system for surgical audit. *Med J Aust* 1988;149:70-4.
27. Lenhard RE Jr, Buchman JP, Achuff SC, Kahane SN, Macmanus CJ. AUTRES — the Johns Hopkins Hospital automated resume. *J Med Syst* 1991;15(3):237-47.
28. Lissauer T, Paterson CM, Simons A, Beard RW. Evaluation of computer generated neonatal discharge summaries. *Arch Dis Child* 1991;66(4 spec no):433-6.
29. Occena LG, Huang LJ. Facilitating discharge summary preparation in small hospitals. *J Soc Health Syst* 1991;3(1):67-90.
30. Christo GG, Marianus BV, Krishnan L, Iyer RS, Venkatesh A. Computerised neonatal case records: a four year experience. *Indian Pediatr* 1992;29(2):173-80.
31. Hohnloser JH, Purner F. PADS (Patient Archiving and Documentation System): a computerized patient record with educational aspects. *Int J Clin Monit Comput* 1992;9(2):71-84.
32. Lowe WW, Ciszek TA, Gallaher KJ. Comprehensive computerized neonatal intensive care unit data system including real-time, computer-generated daily progress notes. *Pediatrics* 1992;89:62-6.
33. Dunn DC, Dale RF. Combined computer generated discharge documents and surgical audit. *BMJ (Clin Res Ed)* 1986;292:816-8.
34. Llewelyn DE, Ewins DL, Horn J, Evans TG, McGregor AM. Computerised updating of clinical summaries: New opportunities for clinical practice and research? *BMJ* 1988;297:1504-6.
35. Smith RP, Holzman GB. The application of a computer data base system to the generation of hospital discharge summaries. *Obstet Gynecol* 1989;73(5 pt 1):803-7.
36. Castleden WM, Stacey MC, Norman PE, Lawrence-Brown MM, Brooks JG. General practitioners' attitudes to computer-generated surgical discharge letters. *Med J Aust* 1992;157:380-2.
37. Adams DC, Bristol JB, Poskitt KR. Surgical discharge summaries: improving the record. *Ann R Coll Surg Engl* 1993;75(2):96-9.
38. Brazy JE, Langkamp DL, Brazy ND, De Luna RF. Do primary care physicians prefer dictated or computer-generated discharge summaries? *Am J Dis Child* 1993;147:986-8.
39. van Walraven C, Rokosh E. What is necessary for high-quality discharge summaries? *Am J Med Qual* (in press).
40. Rand Corporation. *A million random digits with 100,000 normal deviates*. Glencoe (IL): The Free Press; 1955.
41. Meinert CL. *Clinical trials: design, conduct, and analysis*. New York: Oxford University Press; 1986. ch 9.
42. Ware JH, Antman EM. Equivalence trials. *N Engl J Med* 1997;337:1159-61.
43. Makuch R, Simon R. Sample size requirements for evaluating a conservative therapy. *Cancer Treat Rep* 1978;62:1037-40.
44. Spriet A, Beiler D. When can "nonsignificantly different" treatments be considered as "equivalent"? *Br J Pharmacol Ther* 1979;7:623-4.
45. Solomon JK, Maxwell RB, Hopkins AP. Content of a discharge summary from a medical ward: views of general practitioners and hospital doctors. *J R Coll Physicians Lond* 1995;29:307-10.
46. Hart CWM. The Hawthorne experiments. *Can J Econ Polit Sci* 1942;9:150-63.

Reprint requests to: *Dr. Carl van Walraven, Clinical Epidemiology Unit, Loeb Research Institute, Rm. F660, The Ottawa Hospital – Civic Campus, 1053 Carling Ave., Ottawa ON K1Y 4E9; fax 613 761-5351; carlv@lri.ca*

Appendix 1: Laboratory and radiologic data abstracted from each patient's hospital discharge summary, summarative laboratory report and radiologic data summary

Extreme results*	
Hematology	Leukocyte count < 2.0 or > 15.0 x 10^9/L
	Hemoglobin concentration < 100 or > 180 g/L
	International normalized ratio > 5.0
Serum biochemistry	Sodium level < 125 or > 150 mmol/L
	Sodium bicarbonate level < 15 mmol/L
	Creatinine level > 300 μmol/L
	Total calcium level > 3.0 mmol/L
	Creatine kinase level > 200 IU/L with MB isoenzyme fraction > 5%
Microbiology	Blood culture (except for *Streptococcus viridans*)
	Urine culture (> 100 million colony-forming units per litre of urine)
	Cerebrospinal fluid (any organism)
Diagnostic tests†	
Protein-based tests	Thyroid-stimulating hormone
	Parathyroid hormone
	Cholesterol
	Serum vitamin B_{12} (cobalamin)
	Serum or erythrocyte folate
	Ferritin
	Hemoglobin A_{1c}
Serology	HIV
	Hepatitis A
	Hepatitis B
	Hepatitis C
Radiology	Radiography of chest, abdomen or extremity
	Ultrasonagraphy of abdomen or pelvis
	Doppler ultrasonography of carotid arteries or leg veins
	Computed tomography of head, chest, abdomen or pelvis
Nuclear medicine	Ventilation–perfusion scan

*For tests indicating severity of illness.
†Tests helpful for continuing patient care.