Interactive voice response systems for improving delivery of ambulatory care
Oake N, Jennings A, van Walraven C, Forster AJ

CRD summary
This review assessed use of IVRS to improve a number of predefined patient outcomes. There was significant heterogeneity between the included studies and some concern about the quality of the selected studies. The conclusion that this technology can result in modest improvements in adherence to many processes of care should be interpreted with a degree of caution.

Authors' objectives
To comprehensively describe the populations, interventions and outcomes of interactive voice response system (IVRS) clinical trials.

Searching
MEDLINE (1950 to 2008) and EMBASE (1980-2008) were searched. Search terms were reported. Science Citation Index was handsearched. Reference lists of included studies were examined. Studies were restricted to those published in English.

Study selection
Randomised controlled trials (RCTs) and quasi-randomised controlled clinical trials (CCTs) were eligible for inclusion if they examined the effect of an IVRS on at least one of a number of predefined outcomes. Study populations were variable in terms of age and sex. Settings were variable; most of the studies were community based, others were conducted at speciality clinics, health management organisations, home care companies and a Veterans Affairs medical centre. The intervention studied was IVRS (with or without another simple intervention) and the comparator was no IVRS (but included the same simple intervention if this had been used in the intervention group). The most common targeted behaviours studied were immunisation and healthy lifestyle. Some studies were only of patients with diabetes, heart failure or mental illness. Frequency of patient contact with the IVRS was highly variable. Outcomes were grouped into clinical end points (death, hospitalisation), disease control measures (blood pressure, glycosylated haemoglobin, score on validated disease scale), process adherence (adherence to screening tests, immunisation protocols, home glucose monitoring) and quality of life measures (health scores from validated questionnaires).

The authors did not state how the papers were selected for the review.

Assessment of study quality
Studies were allocated a score based on a checklist specified by Cochrane Effective Practice and Organisation of Care group. The checklist included three primary criteria (concealment of allocation, blind assessment of primary outcome and completeness of follow-up) and three secondary criteria (balanced baseline measures, reliable outcome measures and protection against contamination). Studies were considered to be high quality if all three primary criteria were satisfied and there were no significant concerns regarding secondary criteria. Studies were considered moderate quality if one or two of the primary criteria were unclear or not done and low quality if all three primary criteria were unclear or not done.

The authors did not state how the validity assessment was performed.

Data extraction
Data were extracted in order to calculate mean differences or odds ratios (OR) with 95% CIs. Data on outcomes was extracted on an intention to treat basis.

Two reviewers independently abstracted the data.
Methods of synthesis

Because of significant heterogeneity across outcomes, the authors limited the pooling of data to clinical externally validated measures of disease control and dichotomous process adherence outcomes. For measures of disease control, pooled weighted mean differences (WMD) and their 95% confidence intervals (CI) were calculated using a random-effects model. Heterogeneity was assessed using \( \chi^2 \) and \( I^2 \) tests. For dichotomous process adherence outcomes, the overall effect estimate was calculated using median-effects methodology.

Change scores were calculated for objective outcomes that could not be pooled and patient-reported outcomes. Change scores were defined as the difference between the change from baseline to follow-up in the intervention and control groups. This data was tabulated and briefly discussed in the narrative.

Results of the review

Forty studies (n=106,959) were included in the review: 32 RCTs (n=103,682) and eight CCTs (n=3,277). Seven studies were considered to be high quality, 27 studies were regarded as moderate quality and six studies were considered low quality.

Six studies (n=1,009) reported externally validated clinical measures of disease status. IVRS users showed a non-significant improvement in glycosylated haemoglobin, total cholesterol and serum glucose.

Three studies reported clinical end points that could not be pooled. A study on asthma patients showed no improvement in the IVRS group on acute or routine care visits or on hospitalisation for asthma. Two smaller studies on patients with heart failure showed some improvement in number of cardiac events and time in hospital in the IVRS group.

Fourteen studies that measured objective process adherence outcomes were analysed. The median effect of the IVRS intervention was significant at 7.9% (interquartile range 2.8 to 19.5). A subgroup analysis of eight studies of adherence to immunisation advice showed that those who received the IVRS intervention had significantly higher immunisation rates (OR 1.70, 95% CI 1.41, 2.04, \( I^2 = 78\% \)).

Twenty-four studies reported a total of 74 patient-reported outcomes that could not be pooled. Two-thirds of outcomes showed no significant differences between IVRS and control groups and one third had outcomes that favoured the IVRS group.

Five studies reported on quality-of-life outcomes. Two of these studies reported that IVRS was associated with significant improvements in general health.

Authors’ conclusions

IVRS-based interventions were feasible in many settings and can result in modest improvements in adherence to many processes of care. There was insufficient data to show that IVRS-based interventions improved outcomes.

CRD commentary

The review addressed a clear question with well-defined inclusion criteria. Three relevant sources were searched to identify potential studies. No explicit attempts to address language bias or identify unpublished studies was described, which raised the possibility that language and publication biases might have been introduced in the review. The authors attempted to minimise bias and error during the review process by carrying out data abstraction in duplicate. It was unclear how study selection and quality assessment were performed; therefore, reviewer error and bias might have been introduced at these stages. Primary study details were provided, except for cointervention details. The lack of investigation into causes of significant statistical heterogeneity made the reliability and generalisability of the pooled results uncertain. Most studies were described as either moderate or poor quality, which weakened the reliance that could be placed on their findings.

The authors’ relatively conservative conclusion that IVRS interventions can result in modest improvements to adherence to many processes of care should be interpreted with a degree of caution.
Implications of the review for practice and research

Practice: The authors stated that where provider groups were focused on improving specific processes of care, they felt that published data supported the decision to implement IVRS technology. However, they cautioned that there was insufficient data to demonstrate an improvement in outcomes from this technology.

Research: The authors recommended that future studies should evaluate patient outcomes.

Funding
Partly funded by the Ottawa Hospital Centre for Patient Safety.

Bibliographic details

PubMedID
19514804

Original Paper URL

Additional Data URL

Indexing Status
Subject indexing assigned by NLM

MeSH
Ambulatory Care; Humans; Quality Assurance, Health Care; Quality of Health Care; Speech Recognition Software; User-Computer Interface

AccessionNumber
12009107173

Date bibliographic record published
11/11/2009

Date abstract record published
20/04/2011

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.