Systematic review and evaluation of web-accessible tools for management of diabetes and related cardiovascular risk factors by patients and healthcare providers

Catherine H Yu,1,2 Robinder Bahniwal,1 Andreas Laupacis,1,2 Eman Leung,1 Michael S Orr,3 Sharon E Straus1,2

ABSTRACT

Objective To identify and evaluate the effectiveness, clinical usefulness, sustainability, and usability of web-compatible diabetes-related tools.

Data sources Medline, EMBASE, CINAHL, Cochrane Central Register of Controlled Trials, world wide web.

Study selection Studies were included if they described an electronic audiovisual tool used as a means to educate patients, care givers, or clinicians about diabetes management and assessed a psychological, behavioral, or clinical outcome.

Data extraction Study abstraction and evaluation for clinical usefulness, sustainability, and usability were performed by two independent reviewers.

Results Of 12616 citations and 1541 full-text articles reviewed, 57 studies met inclusion criteria. Forty studies used experimental designs (25 randomized controlled trials, one controlled clinical trial, 14 before—after studies), and 17 used observational designs. Methodological quality and ratings for clinical usefulness and sustainability were variable, and there was a high prevalence of usability errors. Tools showed moderate but inconsistent effects on a variety of psychological and clinical outcomes including HbA1c and weight. Meta-regression of adequately reported studies (12 studies, 2731 participants) demonstrated that, although the interventions studied resulted in positive outcomes, this was not moderated by clinical usefulness nor usability.

Limitation This review is limited by the number of accessible tools, exclusion of tools for mobile devices, study quality, and the use of non-validated scales.

Conclusion Few tools were identified that met our criteria for effectiveness, usefulness, sustainability, and usability. Priority areas include identifying strategies to minimize website attrition and enabling patients and clinicians to make informed decisions about website choice by encouraging reporting of website quality indicators.

Diabetes mellitus affects 285 million people worldwide and is a leading cause of death in most high-income countries.1 Clinical care gaps are common in diabetes care. For example, in an American population-based survey, only 62% of patients with diabetes had low-density lipoprotein cholesterol measured annually.2 Reviews of interventions targeting patients and healthcare providers to optimize diabetes care have shown small effects on provider performance and patient outcomes.3–5

Given that consumers are increasingly using the world wide web as a source of health information,6 web-based tools offer potential for optimizing quality of diabetes care. Use of web-based media may improve knowledge, social support, behavior change, and clinical outcomes.7,8 However, existing diabetes websites have wide variations in the quality of evidence provided9 and offer didactic information at high reading levels with little interactive technology, social support, or problem-solving assistance.10 Similarly, although healthcare providers increasingly use online resources for patient care, the volume, breadth, editorial quality, and evidence-based methodology upon which they were developed are highly variable.11

The effectiveness of these tools in changing clinical outcomes has been the subject of reviews in other topic areas; for example, a systematic review of consumer health informatics applications in diverse topic areas, including breast cancer, found that these applications improved clinical outcomes.12 Their effectiveness in a research setting may not translate to effectiveness in clinical practice; factors that affect their adoption into clinical practice include clinical usefulness, usability, and sustainability.13 Specifically, a clinically useful tool, defined as a tool that provides clinically useful answers and is easy to use, access, and read,14 may differ in a research context; for example, while a website on carbohydrate counting may be useful in a research setting with a research dietician, it may be less useful to the consumer trying to use it alone in a real-life setting. Similarly, usability of the tool (defined as the extent to which a product can be used by specified users to complete tasks successfully, in time, and with satisfaction in a specified context15) may be underemphasized in research studies,16 where participants are routinely oriented to and trained on the use of the tool. Finally, sustainability, defined as the degree to which an innovation continues to be used after initial effort to secure adoption is completed,17 is a critical component in addressing the gap between research and practice,18 yet is often not addressed or assessed.19 Critical appraisal of web tools should thus consider their effectiveness and their clinical usefulness, usability, and sustainability. Previous studies have not evaluated the validity, clinical usefulness, usability, and sustainability of web-compatible, diabetes-related tools for patients and providers, which was the objective of this study.

This paper is freely available online under the BMJ Journals unlocked scheme, see http://jamia.bmj.com/site/about/unlocked.xhtml

Because of the importance of multifactorial vascular risk reduction as well as comprehensive lifestyle modification in the care of patients with diabetes, we were interested in diabetes-specific tools and tools for blood pressure, lipid, smoking, obesity, nutrition, physical activity and weight management.

METHODS
Data sources and searches
Published literature search strategy
In consultation with an information scientist, Medline, EMBASE, CINAHL, and the Cochrane Central Register of Controlled Trials were searched from their earliest date to June 1, 2011. The following search terms were used: diabetes, hypertension, smoking cessation, weight reduction, online, computer-based, and internet. The complete search strategy is provided in online appendices. Additional articles were identified through review of reference lists of identified studies and discussions with experts.

Grey literature search strategy
The world wide web was searched using the Google search engine on June 14, 2009 with preselected phrases (online appendices). We used these phrases to search websites of interest that had been identified on the basis of expert knowledge. Sixty web ‘hits’ were captured for every phrase. The first 50 hits that met our definition of a web-compatible diabetes care tool were retained.

Study inclusion and exclusion criteria
Studies were included if they evaluated a web-compatible diabetes-related care tool, defined as an audiovisual tool that is provided in electronic form to be used as a means to educate, support, or advise patients, care givers, or healthcare providers, and that addressed one of the following aspects of diabetes and pre-diabetes management: glycemic control; cardiovascular risk factor assessment and modification (including hypertension, dyslipidaemia, and smoking cessation); nutrition, physical activity, and weight management; self-management and psychological issues; and complication screening and management. Tools were included if they addressed a relevant topic area, regardless of whether patients with diabetes were the target population. Studies could have observational or experimental designs and had to include at least one psychological, behavioral, or clinical outcome.

Studies were excluded if they (1) did not include an evaluation of the tool, (2) were in a language other than English or French, or (3) evaluated a tool that (a) consisted of an electronic database with no exportable stand-alone tool (such as an electronic medical record), (b) had supplementary hardware or software requirements that were not readily available to the average user, or (c) required a real-time operator (such as a healthcare provider).

Study selection
Titles and abstracts were screened for relevance by two independent reviewers (CHY, SES; figure 1). Potentially eligible full articles were then retrieved and reviewed independently by two reviewers (CHY, SES) to determine whether they met inclusion criteria. A third reviewer was available in cases of disagreement.

Determination of tool accessibility
Tools were evaluated for accessibility by attempting to access the web-connected tool on a personal computer with standard software. If the tools were not readily accessible online, two attempts were made to contact the authors by email for additional information or access.

Data extraction and quality assessment
For each study that identified an accessible tool, two reviewers independently extracted study characteristics using electronic data extraction forms. We used a modification of the Cochrane Effective Practice and Organization of Care Group data abstraction form. These forms characterized study design, participants, tool description, study outcomes, and results. Differences in data extraction were reconciled by consensus. Tools identified from the grey literature search were categorized by content and educational focus (online appendices). We randomly selected five websites from each of these categories and reviewed them to determine if there was evidence of clinical effectiveness. To assess these sites, we developed an instrument based on a framework by Straus and Haynes and tested its face validity with relevant experts. This instrument contained 18 items and characterized the evidence base for the content and effectiveness of the tool.

Tool evaluation
Clinical usefulness and sustainability
Each tool was independently reviewed for clinical usefulness and sustainability by two members of a team of clinical experts (CHY, F Kim, H Halapy, and C West; see Acknowledgments). Differences were reconciled by consensus. As there were no validated instruments to assess clinical usefulness or sustainability, we developed instruments using a framework from the literature and input from experts; these instruments were determined to have face validity by experienced clinicians and experts in knowledge translation. The clinical usefulness instrument contained five items and assessed clinical relevance and ease of access using a Likert scale with scores ranging from 0 to 5. A score of 5 denoted ‘clinically useful answers are available most of the time, and are easily accessible and readable within a few minutes’, and a score of 0 denoted ‘not useful clinically.’ The sustainability instrument contained six items and assessed continued relevance of the topic, potential barriers to sustainability, and engagement of a group to keep the tool up to date. The instrument was designed to exclude major barriers to sustainability.

Usability
Each tool was independently reviewed for usability by two members of a team of human factors engineers (S Jovicic, A Xu, H Takeshita, and F Wan; see Acknowledgments). The instrument incorporated questions from three industry-standard usability instruments (ISO 9241-110 Usability Heuristics, Gerhardt-Powals Research-based Heuristic, and Site Assessment Tool) and contained 27 items characterizing suitability to user’s skill, ease of navigation, reduction in cognitive load, and appearance. For each desirable usability characteristic, raters scored ‘yes’, ‘no’, or ‘not applicable.’

Data synthesis and analysis
Inter-rater reliability for data abstraction for clinical effectiveness, clinical usefulness, sustainability, and usability were calculated.

Owing to heterogeneity in study design, population, interventions, and outcomes, meta-analyses by intervention type or outcome were not possible. However, we described study quality and performed a descriptive analysis of studies with evidence of impact on outcomes. In addition, we performed a meta-regression of all eligible studies (irrespective of intervention
type or outcome) to assess whether clinical usefulness ratings and usability ratings were moderators of the effectiveness of these interventions. The meta-regression was performed in R²⁷ version 2.12.0 using the contributed package metafor²⁸ version 1.4-0. The Mad²⁹ version 0.8 package was used to convert the treatment effects into a standardized treatment effect. Hedges G³⁰ Studies were excluded if there were insufficient data to determine a treatment effect or its variance. Studies without a ‘true’ control group were also excluded, as there is no way to incorporate clinical usefulness or usability for both groups.

RESULTS
Published literature
Results of the literature search, study, and tool selection are detailed in figure 1. While 395 studies and 219 unique tools were identified, 127 tools were not accessible, and thus we were not able to evaluate them. We evaluated the remaining 92 tools and corresponding 57 studies.

Inter-rater reliability
Inter-rater reliability was moderate to good: κ for data abstraction items for clinical effectiveness ranged from 0.66 to 0.72. Weighted κ values³¹ for assessment of clinical usefulness and sustainability were 0.50 (95% CI 0.36 to 0.64) and 0.77 (95% CI 0.63 to 0.87), respectively. Cohen’s κ³² ranged from 0.50 to 0.65 for each component of the scale (ISO, 0.65; Gerhardt-Powals, 0.50; Site Assessment Tool, 0.55).

Description of studies
Study quality and type
Of 57 studies, 40 studies used experimental designs (25 randomized controlled trials (RCTs),³³—³⁵ one controlled clinical trial,³⁶ 14 uncontrolled before—after studies,³⁷—³⁸ and 17 studies used observational designs (one case—control trial,³⁹ seven cross-sectional studies,⁴⁰—⁴⁶ nine cohort studies⁴⁷—⁵⁰). One article consisted of four studies including two RCTs and two uncontrolled before—after studies.⁴⁷ One article consisted of one RCT and one uncontrolled before—after study.⁴⁸

Risk of bias
The methodological quality of all studies is described in online appendix tables 1 and 2. Characteristics of the RCTs are summarized in appendix table 1. Methodological quality was variable; intention-to-treat analysis and description of loss to
follow-up occurred in approximately half of the studies, and calculation of statistical power, randomization, concealment of allocation, and follow-up of more than 6 months were described in the minority of studies.

**Description of tools**

**Tool formats**

Formats included static websites, decision aids, interactive websites, CD-ROM games or DVD, and email feedback programs.

**Target audience**

Five tools targeted patients with type 1 or 2 diabetes. The remaining tools did not specifically target patients with diabetes, but did address an aspect of comprehensive diabetes care in overweight adults, smoking adults, depressed adults, children, smoking adolescents, and adolescents at risk of type 2 diabetes. With respect to targeted care givers and healthcare professionals, 10 studies targeted healthcare providers, with six targeting physicians, and one study targeting public health professionals; there were no studies that targeted care givers.

**Clinical usefulness, sustainability, and usability**

Clinical usefulness, sustainability, and usability ratings are summarized in table 1. The most common usability error (found in 50% of tools) was not utilizing images to facilitate learning, a feature that has been demonstrated to aid data interpretation and improve recognition and recall.

**Clinical effectiveness of tools**

Patient outcomes, including knowledge, skill development, behavior change, and psychological and clinical outcomes, were examined in 17 studies. Clinician knowledge and skill were evaluated in three studies (online appendix table 3).

**Patient diabetes education tools**

One study examined the effect of a multimedia general diabetes education computer application for low-literacy patients to use in clinic waiting rooms. At 1 year, the intervention group had a greater awareness of diabetes complications and a greater reduction in HbA1c than the control group (online appendix table 5).

**Patient self-management and coping tools**

Two studies examined the effect of a self-management website or DVD in patients with poorly controlled diabetes. In the first study, the intervention group had a greater reduction in HbA1c and systolic blood pressure, an increase in high-density lipoprotein cholesterol, and reduction in triglycerides compared with the control group at 12 months (online appendix table 5). In addition, greater website use correlated with greater clinical improvements: persistent website users had greater reduction in HbA1c from baseline compared with intermittent users (−1.9% vs −1.2%, p=0.051). Similarly, larger numbers of website data uploads were associated with a larger decline in HbA1c (highest

---

**Table 1** Summary of clinical usefulness, sustainability, and usability ratings

<table>
<thead>
<tr>
<th>Summary of clinical usefulness ratings</th>
<th>Number of tools for which:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Had clinically useful answers available:</td>
<td>Answers were easily accessible within a few minutes</td>
</tr>
<tr>
<td>Most of the time</td>
<td>26</td>
</tr>
<tr>
<td>Some of the time</td>
<td>18</td>
</tr>
<tr>
<td>Rarely</td>
<td>6</td>
</tr>
<tr>
<td>Not at all</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Summary of sustainability ratings</th>
<th>Number of tools with response of:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sustainability instrument item</td>
<td>Yes</td>
</tr>
<tr>
<td>Will this topic continue to be relevant?</td>
<td>32</td>
</tr>
<tr>
<td>Are there any potential barriers for patients, care givers, the public, or healthcare providers to using this tool?</td>
<td>8</td>
</tr>
<tr>
<td>Can this tool be easily integrated into existing practice and systems?</td>
<td>23</td>
</tr>
<tr>
<td>Can groups be easily engaged to facilitate sustainability of this tool?</td>
<td>26</td>
</tr>
<tr>
<td>Is there a leader responsible for making modifications to this tool as new knowledge is brought forward?</td>
<td>29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Summary of usability ratings</th>
<th>Number of tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of usability errors</td>
<td>References</td>
</tr>
<tr>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>1−2</td>
<td>31</td>
</tr>
<tr>
<td>3−5</td>
<td>35</td>
</tr>
<tr>
<td>6−10</td>
<td>21</td>
</tr>
</tbody>
</table>
tertile –2.1% vs lowest tertile –1.0%, p<0.02). The second study, which compared in-person class-based programs with DVD-based self-management programs, showed no differences in self-management behavior or psychological or clinical outcomes.33

One study examined the effect of an interactive website providing tailored advice on lifestyle modification and risk factor screening, based on a questionnaire on family history and lifestyle habits.35 Compared with a control group who received standardized messages, the intervention group improved their diet and physical activity, although there were no changes in smoking rates and assessment of blood glucose or blood pressure, and a reduction in cholesterol screening.

**Patient nutrition and physical activity tools**

Seven studies examined the effect of nutrition and physical activity websites37 40 44–46 52 53 on waist circumference,40 weight,53 body mass index (BMI),44 percentage body fat,44 blood pressure,44 quality of life,37 and coronary RR.40 Measures of obesity (waist circumference, weight, BMI, and percentage body fat) were significantly improved,40 44 53 as was coronary RR 44 and self-assessed health status37 (online appendix table 5). In a study of a web-based physical activity program,40 waist circumference decreased significantly in the intervention group compared with the controls. Similarly, an interactive personalized health promotion website was found to result in a greater reduction in percentage body fat compared with the control, although there was no significant change in BMI or blood pressure.44 This study also found a significant reduction in coronary RR in the intervention group, compared with the control group. A third study compared the effect of a weight loss website in combination with human-email counseling, computer-automated email counseling, or no counseling46 and found that, at 3 months, there was a significantly greater weight loss in the human-email group and computer-automated email group than in the no counseling group. However, at 6 months, only the human-email group retained significantly greater weight loss compared with the no counseling group. Finally, self-assessed health status was significantly greater in the group using an email physical activity and diet program than the control group.57

One study examined the effect of an educational CD-ROM game about nutrition and physical activity for children.46 Whereas there was a greater reduction in BMI in girls (p=0.04), a greater increase in BMI was noted in boys (p=0.04) 5 weeks after the intervention.

**Patient smoking cessation tools**

Eight studies examined the effect of five online smoking prevention and cessation websites on clinical outcomes (smoking initiation,50 cigarette use,48 1-day smoking abstinence,55 7-day reported abstinence,54 47 48 30-day point prevalence,51 56), and one study showed an improvement in quit rate.47 However, in an exploratory analysis by website utilization, Rabius found that higher smoking-quit rates were associated with the two more highly utilized websites compared with the three less frequently utilized sites (12.5% vs 10.6%, p=0.03).51 Similarly, participants who visited a site more than five times were twice as likely to quit than participants who visited a site less than five times (20.0% vs 9.8%, p<0.001). In addition, higher quit rates were found with more interactive, tailored sites compared with the static control site (15% vs 10%, p=0.04).51

**Clinician education tool**

One study examined the effect of an online continuing education seminar on physician knowledge of diabetes management.54 Physicians’ recommendation of appropriate quality-of-care measures was assessed immediately after the intervention using a non-validated clinical vignette score and did not change, with the exception of one process-of-care measure (‘ordering an eye exam’).

**Clinician behavior change counseling tool**

Two studies examined the effect of an interactive web-based motivational interviewing educational program on teaching effectiveness immediately after the intervention.38 39 In both studies, teaching effectiveness, as measured by qualitative analysis and coding of written responses to counseling vignettes and a multiple choice questionnaire, was higher in the intervention group.

**Role of interactivity**

More interactive tools resulted in greater clinical improvement; for example, Tate et al found that interaction in the form of human- or computer-email counseling resulted in greater weight loss than no counseling.53 Similarly, Coran and Reynolds found that an interactive multimedia CD-ROM game resulted in greater reduction in BMI than a static educational CD-ROM.43 This observation is seen also in patient smoking cessation websites; Rabius et al found that interactive tailored smoking cessation sites resulted in greater quit rates than a static site,51 and Munoz et al found that individually timed educational messages resulted in greater quit rates than the static smoking guide alone.47 A similar finding was seen in tools for healthcare providers; Carpenter et al found that an interactive tutorial was more effective in teaching motivational interviewing techniques than reading material.38 39 The role of tool interactivity on continued website use is highlighted in the study of the interactive personalized health promotion website: Hurling et al found that the interactive site resulted in a significantly greater percentage of participants logging in throughout the study period, with less attrition than the static, control site.45

**Grey literature**

Tool selection and evaluation are described in figure 2. Of the 360 websites reviewed, two56 87 had been evaluated for clinical effectiveness. Both evaluations had been identified in the published literature search47 48 and the tools described previously.

**Clinical usefulness and usability as potential moderators of tool effectiveness**

Figure 3A shows the results of the meta-regression, with clinical usefulness ratings and usability ratings as potential moderators of tool effectiveness. Twelve studies were included, comprising a total sample size of 2751 participants. There was
significant heterogeneity, with $\tau$ of 0.87 and 0.85, respectively. While the standard meta-analysis demonstrated a significant positive effect on outcomes (standardized treatment effect, Hedges $G$ 0.64, 95% CI 0.15 to 1.13, $p=0.01$), neither clinical usefulness nor usability had a moderating effect on tool effectiveness (regression coefficient 0.26 (95% CI $-1.4$ to $1.9$, $p=0.76$) and $-1.5$ (95% CI $-6.4$ to $3.4$, $p=0.55$), respectively).

**DISCUSSION**

Although a large number of studies and tools were identified, many tools were not accessible, and thus we were not able to evaluate them. These tools would also not be accessible to patients or healthcare providers; thus their exclusion does not affect the applicability or relevance of our findings. The 57 studies and tools identified were very different in terms of participants, settings, and outcomes, which meant we could not perform a meta-analysis by intervention or outcome. Although there were a number of studies with positive results, these results must be viewed with caution because of concerns about the reported study designs. Half of the studies were pre—post designs or included a comparative group that was non-randomized or not adequately randomized. Many studies would have been strengthened through use of validated outcome measures and longer-term follow-up of 1 year or more. With respect to the tools, although the evidence base of the tool’s content was high, only 25% had easily accessible, clinically useful answers most of the time. Six percent of tools were free of usability errors, but 60% had three or more errors in usability. Common usability errors included limited use of images, icons, and other visual elements to facilitate learning, and lack of intuitiveness in navigation and expected next steps. These and other usability errors can negatively affect users’ experience with a tool and may lead them to stop using the tool. In one study, ease of usability was one of the main determinants of an individual’s satisfaction and willingness to engage with a website.88 While a correlation between usability and tool effectiveness was not demonstrated in this study, our meta-regression was limited by the number of studies that adequately reported data, as well as the heterogeneity in interventions, populations, and outcomes. This high prevalence of usability errors is mirrored in other reviews of usability of healthcare websites89—91 and highlights the need to ensure that websites provide useful and usable formats and undergo usability testing before they are launched.

Our review of the literature has identified areas for further exploration. First, greater improvements in patient outcomes were seen with greater use of the tool.42 49 51 55 For example, persistent website users had greater improvement in HbA1c than intermittent users, and a larger number of website data uploads was associated with a larger decline in HbA1c.55

![Figure 2](image-url)  
**Figure 2**  Grey literature search algorithm. Modified PRISMA flow diagram outlining results of grey literature search and tool identification process.

![Figure 3](image-url)  
**Figure 3**  (A) Modified forest plot demonstrating lack of moderating effect of clinical usefulness ratings on tool effectiveness. Squares with lines are the observed treatment effects and CI. Grey diamonds show the predicted treatment effects based on the model. (B) Modified forest plot demonstrating lack of moderating effect of usability ratings on tool effectiveness. Squares with lines are the observed treatment effects and CI. Grey diamonds show the predicted treatment effects based on the model.
Caution should be used when interpreting this association, given possible confounders that can result in reduced use such as depression or lack of progress with respect to goals. A recent systematic review found mixed results in the association between adherence and outcomes; analysis was limited because of heterogeneity of adherence and outcome measures, although logins appeared to be associated with outcomes of physical health interventions, while module completion appeared to be associated with outcomes of psychological health interventions. In conjunction with the literature on website usage attrition, these findings have implications for website development and website evaluation. Second, this review suggests a mechanism by which to minimize attrition and thus maximize clinical improvement, through the use of interactivity and feedback. For example, Hurling et al found that an interactive health promotion site resulted in a significantly greater percentage of participants logging in throughout the study period, with less attrition, than the control static site. Although greater interactivity may result in better outcomes, it may require higher levels of health literacy, navigation skills, and computer experience. However, with careful user testing, highly interactive applications can be designed to be user friendly and can have positive effects on user satisfaction, effectiveness, efficiency, and overall attitude toward the tool. Other intervention characteristics that enhance use include peer or counselor support, email or phone contact, and updates regarding the intervention website.

These findings have implications for website developers, researchers, patients, and clinicians. Web-based tool developers must incorporate strategies—such as optimization of website usability and interactivity—to maximize frequency and persistence of website use, and researchers must evaluate these strategies and their impact on website usage and clinical outcomes, as well as characteristics of users who are predisposed to persistent website use. Given the degree of variability in website quality, patients and clinicians should critically appraise these resources for effectiveness, relevance, and usability before selecting them for use. Given the burden of time and expertise required to make these assessments, development of a transparent recognized peer-review system to assess clinical effectiveness, usefulness, sustainability, and usability of web-based tools, as well as a requirement for standard reporting of these characteristics by website developers, would enable both patients and clinicians to make informed decisions in a timely manner. Although website certification systems do exist, these primarily address the evidence base of the website content rather than website use, and do not address clinical usefulness or usability. For example, Health On the Net Foundation (HON, http://www.hon.ch/) is a non-profit organization that established the HONcode certification, an ethical standard aimed at offering quality online health information. A review of HONcode-accredited sites found that 87% were too difficult to read for the average adult population. In addition, this system has not been universally adopted, with only 28% of diabetes patient education sites being HONcode-accredited. This review is limited by the number of accessible tools, exclusion of tools for mobile devices, the quality of the studies identified, use of non-validated scales, and publication bias. In addition, the broad interventions included, as well as the number of study outcomes (clinical effectiveness, clinical usefulness, sustainability, and usability), limited the ability to synthesize data with a standard meta-analytical approach. Mobile devices represent a highly accessible portal to health information resources and thus have the potential to transform healthcare delivery; however, assessing tools for mobile devices was beyond the scope of this review. Although the rating scales used were not formally validated, the items were derived from the literature and were assessed for face validity by content experts. We chose to be inclusive when selecting interventions, given the multi-system involvement of the diabetic disease process and the importance of comprehensive management (including vascular risk modification) in the care of the individual with diabetes. We also chose to assess the non-traditional outcomes of clinical usefulness, sustainability, and usability, as these are important predictors of knowledge use and transfer. We strove to reduce publication bias by including a comprehensive search of the grey literature.

The strengths of this review include: an extensive literature search that included the grey literature; the comprehensive review of each tool for clinical effectiveness, usefulness, sustainability, and usability; and the generalizability of findings regarding website use in health promotion and chronic disease management. To our knowledge, although other reviews of health informatics tools have addressed clinical outcomes and usability individually, no other systematic review of any informatics intervention has considered all of these issues.

Web-based tools have the potential to improve health outcomes and complement healthcare delivery, but their full potential is hindered by limited knowledge about their effectiveness, high prevalence of usability errors, and high attrition rates. A development and research agenda should include: developing strategies to reduce website attrition in order to maximize clinical outcomes; standardizing website quality indicators; and transparent reporting of these indicators in order to allow patients and clinicians to make informed decisions about website choice.

Acknowledgments We thank: Laure Perrier for carrying out the published and grey literature searches; David Newton for information technology support; Julie Huki, Daniel Chong, Patrick Ong, and Kristin Tokarsky for reference management; Mahmood Beheshti, Nathan McKibbon, Alex Rogovic, and Violetta Sochka for website review; our clinician experts (Florence Kim, Henry Halapy, and Christine West) for rating clinical usefulness and sustainability; and our human factors engineers (Sasha Jovicic, Annie Xu, Harumi Takeshita, and Flora Wan) for rating usability.

Funding This research was supported by the Ontario Ministry of Health and Long Term Care, who had no role in study design, in the collection, analysis and interpretation of data, in the writing of the report, or in the decision to submit the article for publication. SES is supported by a Tier 1 Canada Research Chair.

Competing interests None.

Contributors CHY conceived the study, participated in its design, and drafted the manuscript. EL and MSO performed data abstraction and interpretation of data, in the writing of the report, or in the decision to submit the article for publication. RB participated in the coordination, performed data abstraction, and drafted portions of the manuscript. SES and AL participated in its design and revised the manuscript critically for important intellectual content. All authors had access to the data and read and approved the final manuscript.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES
