Operationalization of Skin Self-Examination in Randomized Controlled Trials with Individuals at Increased Risk for Melanoma: A Systematic Review

Adina Coroiu¹-², Chelsea Moran³, Catherine Bergeron², Brett D. Thombs²-⁴-⁸, Alan C. Geller¹, Emily Kingsland⁹, Annett Körner²-⁴,¹⁰-¹¹

Authors’ Affiliations

¹ Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States of America
² Department of Educational and Counselling Psychology, McGill University, Montréal, Quebec, Canada
³ Department of Psychology, University of Calgary, Calgary, Alberta, Canada
⁴ Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Quebec, Canada
⁵ Department of Psychiatry, McGill University, Montréal, Quebec, Canada
⁶ Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, Quebec, Canada
⁷ Department of Medicine, McGill University, Montréal, Quebec, Canada
⁸ Department of Psychology, McGill University, Montréal, Quebec, Canada
⁹ McGill Library and Archives, McGill University, Montréal, Canada
¹⁰ Louise Granofsky Psychosocial Oncology Program, Segal Cancer Center, Montreal
¹¹ Psychosocial Oncology Program, McGill University Health Centre, Montreal

Corresponding author:
Adina Coroiu, PhD, Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States of America
Email: acoroiu@hsph.harvard.edu
**Funding statement**
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Dr. Coroiu’s research training was supported by graduate and postgraduate awards from the Canadian Institutes of Health Research (CIHR) and the Fonds de Recherche du Quebec-Santé (FRQS). Dr. Thombs was supported by a FRQS researcher salary award.

**Conflict of Interest**
Alan C. Geller was the senior author of the Geller et al., 2006 trial, which is one of the trials included in this review. All other authors declared that they do not have any conflict of interest.

**Authors’ Contribution:**
AC, AK, BT, AG contributed to the design of the study.
EK designed and conducted the database search.
AC, CM, CB conducted the study selection and data extraction.
AC conducted the data synthesis and completed the first draft of the manuscript.
CM, CB, BT, AG, EK, AK Contributed critical feedback to earlier drafts of the manuscript, reviewed, and approved the final draft of the manuscript.
Abstract

Objective

To investigate how skin self-examination was operationalized and the psychometric properties of the scales used to assess this behavior in randomized controlled trials (RCTs) testing interventions that promote SSE among individuals at increased risk for melanoma.

Method

Eight scientific databases (e.g., Medline, EMBASE, CINAHL, PsycINFO) and four trial registries (e.g., Clinicaltrials.gov, UK Clinical Trials Gateway) were searched from inception through April 2, 2019. Three reviewers carried out the selection of relevant trials and conducted data extraction.

Results

The review identified 13 unique RCT’s. The definition of skin self-exams, extrapolated from instructions provided to participants during the trials and reported in only 6/13 trials, included periodically checking the skin of the entire body, individually or with partners/mirrors, with or without tracking or monitoring tools, and using the ABCDE criteria to identify early signs of melanoma. There was variability in how skin self-examination behavior was measured with respect to item content, number of items, response format, and type of outcome variable used: continuous or binary). No validity evidence and minimal reliability evidence for the measures were identified.

Conclusions and Practice Implications

Future studies are needed to establish the psychometric properties of measures assessing skin self-examination.

Registration

PROSPERO (CRD42016033765).

Word count: 197

Keywords: skin self-examination (SSE); validity; melanoma prevention; systematic review; randomized controlled trials (RCT)
1. Introduction

Skin cancers are the most frequently diagnosed cancers in the United States and Canada (Rogers, Weinstock, Feldman, & Coldiron, 2015). Melanoma, which grows and spreads faster than non-melanoma skin cancers (Safarians, Sternlight, Freiman, Huaman, & Barsky, 1996; Yang et al., 2009), is the most lethal type of skin cancer (American Cancer Society, 2016; National Center for Health Statistics, 2011; National Cancer Institute, 2011). Survival is excellent when the disease is detected at an early stage and treated (> 98%), but considerably lower when the disease has spread to other organs (23%) (Howlader et al., 2016). Tumour thickness/depth of invasion at diagnosis is the best predictor of survival (Baade et al., 2006; Balch et al., 2009; Eisemann et al., 2012; Green, Baade, Coory, Aitken, & Smithers, 2012). Thus, early detection and timely treatment (surgical excision of the tumour) are crucial to survival. Because many melanomas develop with a visible pre-clinical phase (i.e., visible changes on the skin) they are amenable to early detection by means of visually inspecting the skin by physicians and lay persons (Friedman, Rigel, & Kopf, 1985; Hamidi, Peng, & Cockburn, 2010; Weinstock, 2000). Clinical skin exams performed by trained physicians are more accurate at detecting melanoma compared to self-exams performed by lay persons (De Giorgi et al., 2012; Pollitt et al., 2009; Swetter, Pollitt, Johnson, Brooks, & Geller, 2012). However, relying entirely on clinical skin exams to diagnose melanoma early is not a realistic strategy for the majority of individuals living in North America, where fewer than 20% of adults receive annual preventive health check-ups (Mehrotra, Zaslavsky, & Ayanian, 2007). Given that the vast majority of melanomas (50-80%) are detected by patients and family or other lay people (Carli et al., 2003; De Giorgi et al., 2012; Swetter et al., 2012), clinical care guidelines recommend that individuals with an increased risk of developing melanoma, including melanoma survivors (Youlden, Youl, Soyer, Aitken, & Baade, 2014), their first-degree relatives, individuals with many moles, and those with certain phenotypic features, such as blond or red hair color and a tendency to freckle (Gandini et al., 2005), perform regular skin self-examinations in between medical appointments (Marciano, Merlin, Bessen, & Street, 2014; Watts et al., 2015).

Common visual inspection methods used for melanoma early detection include the ABCDE criteria, the 7-point Glasgow checklist, and the “ugly duckling” method, which have been developed for clinicians and the general public (Tsao et al., 2015). The ABCDE criteria (Abbasi et al., 2004; Friedman et al., 1985) target the identification of problematic moles on account of Asymmetry, irregular Borders, varying shades and Colours, large Diameters (> 6 mm), and whether they are Evolving or changing over time. The 7-point Glasgow checklist (MacKie et al., 2007) consists of 7 key features more commonly associated with melanoma than with other skin cancers (sensory change, diameter [≥1 cm], lesion growth, irregular edge, irregular pigmentation, inflammation, and crusting, oozing, or bleeding), which were further refined to include 3 primary (change in size, shape or color) and 4 secondary features (inflammation, crusting or bleeding, sensory change, and diameter [≥ 7 mm]) (Healsmith, Bourke, Osborne, & Graham-Brown, 1994). The “ugly duckling” strategy (Grob & Bonerandi,
aims to identify moles or lesions (skin spots) that stand out when compared to other moles present on the skin in terms of appearance and / or sensations.

Melanoma prevention and advocacy agencies (e.g., American Academy of Dermatology, Canadian Cancer Society, Cancer Care Ontario in Canada), and national policy makers (Macbeth, Newton-Bishop, O’Connell, & Hawkins, 2015; National Collaborating Centre for Cancer, 2015) recommend that individuals with an increased risk for melanoma compared to the general population receive yearly whole-body skin exams from physicians and be advised to perform skin self-exams (SSE) in between medical appointments and seek medical advice for self-identified problematic lesions. However, recommendations on how to perform SSE and the recommended frequency of SSE are not consistent across recommendations for individuals in similar risk categories (Watts et al., 2015) and do not necessarily map onto specific, well-defined criteria (e.g., ugly duckling). For example, the most recent NICE guidelines for melanoma follow-up in the public healthcare system in the UK recommend the delivery of both oral and written instructions on the importance of SSE to patients and their families, but fail to specify how to actually perform a SSE and what strategy to use to identify problematic nevi (NICE, 2015). In addition, there is variability in the recommended frequency for SSE across various guidelines, including monthly SSE, SSE every 3-6 months, or interval not stated (Marciano, Merlin, Bessen, & Street, 2014; Watts et al., 2015). Inconsistent or vague recommendations about SSE can have negative implications for research (e.g., difficulty to assess the impact of adherence to the medical recommendations using meta-analytic procedures) and clinical practice (e.g., if clinicians cannot provide concrete recommendations about how SSE should be performed, it would likely negatively impact the implementation and uptake of SSE behaviors by patients). To date there has been no research study to investigate and synthesize how SSE behaviour is operationalized in melanoma prevention intervention research.

1.1. Research Objectives

The objective of this systematic review was to investigate how skin self-examination (SSE) was operationalized (defined and measured) and the psychometric properties of the measures used to assess SSE in randomized controlled trials (RCTs) testing interventions that promote SSE among individuals at an increased risk for melanoma. Specifically, we addressed the following research questions:

1. How was SSE defined throughout the recommendations provided to participants and did these recommendations map onto existing criteria for the identification of problematic moles?

2. How was SSE measured and was there consistency in the assessment methods across trials?

3. Was there evidence for the validity and reliability of the SSE measures used across trials?

2. Method

The research questions addressed in this review pertain to the secondary objective of a larger systematic review, which aimed to assess the effect of behavioral interventions with individuals at increased risk for melanoma on health and
health-related behaviors, including SSE. This systematic review was registered with the International Prospective Register of Systematic Reviews (CRD42016033765).

2.1 Search Strategy

The search strategy was developed by a research librarian (EK) in collaboration with the research team and was peer-reviewed by a second librarian from McGill. Medline, EMBASE, CINAHL, PsycINFO, Web of Science (Conference Proceedings), ProQuest Dissertations and Theses Global, PubMed, and the Cochrane Central Register of Controlled Trials were searched from inception of each database to April 2, 2019. The following trial registries were also searched: Clinicaltrials.gov, UK Clinical Trials Gateway, International Clinical Trials Registry Platform Search Portal, and the Australian and New Zealand Clinical Trials Registry. The search strategy (see Appendix A) was initially developed for Medline and adapted to the remaining databases. Results from all databases were imported into EndNote, where duplicate removal was performed.

2.2 Identification of Eligible Studies

2.2.1. Study design. Randomized controlled trials (RCTs). Studies with any other study design were excluded.

2.2.2. Population. Adults (18 years of age or older) at increased risk for developing melanoma, as defined by primary studies (e.g., patients with a personal history of melanoma; first degree relatives of melanoma patients; immunosuppressed individuals; survivors of childhood cancers treated with radiation; individuals with phenotypic features such as freckles, fair skin or hair color, presence of more than 5 nevi). For studies that included a mixed population of eligible and non-eligible individuals, we included the study if the eligible individuals could be differentiated from the non-eligible individuals and if statistics were provided separately for our population of interest. For a detailed list of risk categories and eligibility criteria, please refer to Appendix B.

The current review focused on populations with an increased risk of melanoma incidence, primarily because current North American clinical practice guidelines recommend SSE as a melanoma prevention strategy for these populations, but not for the general population. We therefore excluded studies conducted with a general population, populations without an elevated risk, including White men > 50 years of age, who are at higher risk of dying from melanoma than any other group, but do not have a higher incidence risk compared to the general population. Notably, most intervention studies conducted with the general population and men > 50 years of age were conducted in Australia, which has the highest incidence of melanoma in the world. This review aims to inform public health policy in Canada and the United States, where melanoma incidence is significantly lower than in Australia and the clinical guidelines consistently recommend SSE only for populations with a higher risk of incidence compared to the general population.

2.2.3. Interventions. Eligible interventions included informational (passive delivery of information on prevention or early detection), behavioral (active demonstration of preventive behaviors or skills), psychological (delivery of counselling or
psychotherapeutic interventions to assist with the adoption of preventive behaviors), and clinical strategies (non-pharmacological or non-surgical interventions that can only be delivered by medical professionals: GP and dermatologists; clinical skin exams, genetic counselling). Interventions could be delivered in any format (e.g., in person, via pamphlets, via the internet) and administered by professionally trained (e.g., nurses) or non-professionally trained individuals (e.g., research assistants, peers).

2.2.4. Comparison. We included RCTs with any non-active control (e.g., standard of care, treatment as usual, attention control, no treatment) or active control (e.g., another active intervention).

2.2.5. Outcomes

2.2.5.1. SSE conceptualization. This refers to the conceptual explanation of what a self-exam entails, as described in the trial registration and/or the trial report, specifically the section describing the content of the interventions and the instructions and recommendations about SSE given to participants.

2.2.5.2. SSE measurement and consistency of SSE methods. This pertains to the instruments used to assess SSE behavior across trials, including the specific items or scales, the response options, the scoring (how the construct/variable was used in analyses), and the timeframe of assessment (the time span for which the SSE behavior was assessed). We also investigated whether the SSE assessment methods were used consistently across trials.

2.2.5.3. Reliability and validity of the SSE instruments. The investigation of the validity and reliability of the SSE instruments used in trials was conducted with respect to several domains included in the COSMIN checklist (COSMIN; Mokkink et al., 2010) checklist: content validity (the degree to which the content of the instrument accurately reflects the construct that it purports to measure), criterion validity (the degree to which the scores of an instrument adequately reflect a "gold standard" of assessment), and construct validity (the degree to which the scores of an instrument are consistent with hypotheses based on the assumption that the instrument accurately measures a specific construct); internal consistency reliability (the degree of interrelatedness among items) and test-retest reliability.

2.3. Study Selection

Three raters (AC, CM, CB), two in parallel, reviewed the titles and abstracts of all identified citations using a predetermined coding manual (see Appendix C) using Distiller SR (Evidence Partners, 2017), a specialized software for systematic reviews. If either one of the raters deemed an abstract as potentially eligible, a full review was undertaken, also done independently by two raters. Disagreements after the full text review were resolved by consensus between the raters, with the possibility of a fourth team member (AK) to aid with the decision making.

2.4. Data Extraction

Three raters (AC, CM, CB), two in parallel, independently extracted and entered data from relevant trials into pre-designed excel spreadsheets. Data extracted included study characteristics (first author's last name, publication year, country
of study, funding source, trial registration number, population, type of intervention and control, N randomized to intervention and control, age and gender descriptive statistics) and data pertaining to the research questions. For outcome 1, SSE conceptualization, we extracted information about a) the instructions given to patients about how to conduct SSE (how to perform SSE), and b) the recommended frequency of SSE (how often to perform SSE). For outcome 2, SSE measurement and consistency of assessment methods, we extracted information pertaining to the items inquiring about SSE behavior, the answer choices, and the scoring used in data analyses. Based on these data, we drew conclusions about differences and similarities in methods used for SSE assessment across trials. For outcome 3, validity and reliability of SSE instruments, we extracted validity and/or reliability data from the trial report and from the additional sources cited in the method sections of the trial reports, which we identified through manual searching of reference lists. Additional referenced works potentially relevant for the validity or reliability of the SSE instruments used in the included trials were deemed eligible if a) the aim of the study was to evaluate the properties of an SSE instrument identical or similar to the one used in our included trials and/or b) the study reported results for an SSE instrument (identical or similar to one used in our included trials) in the context of investigating its psychometric properties. Disagreements about data extraction were resolved by consensus, with a senior investigator (AK) consulted when necessary.

2.5. Data Presentation and Synthesis

Results were presented in tabular form and a narrative synthesis was provided for each research question.

3. Results

3.1. Results of the Search

The PRISMA flow chart reflecting the study selection process was included in Figure 1. The combined search of the databases generated 667 unique citations, of which 571 were excluded after reviewing titles and abstracts and 75 after the full-text review; 13 unique trials were included.

Through manual searching of the method sections and reference lists of the trial reports included in this review, we identified 18 additional articles potentially relevant to the validity and/or reliability of the SSE measures used in the trials. Upon screening of the full texts (n=18, included in Appendix D), none met our a priori validity and/or reliability criteria for inclusion.

3.2. Characteristics of the Included Trials

Detailed sample characteristics of the included trials can be found in Table 1. Most trials (11/13) were conducted in the United States (Bowen, Burke, Hay, Meischke, & Harris, 2015; Geller et al., 2006; Glanz, Schoenfeld, & Steffen, 2010; Glanz et al., 2015; Glanz et al., 2013; Heckman, Darlow, Ritterband, Handorf, & Manne, 2016; Manne et al., 2010; Oliveria et al., 2004; Robinson, Turrisi, Mallett, Stapleton, & Pion, 2010; Robinson, Turrisi, & Stapleton, 2007a; Turrisi, Hultgren, Mallett, Martini, & Robinson, 2015); 1 was conducted in the United Kingdom (Glazebrook, Garrud, Avery,
Coupland, & Williams, 2006) and 1 in France (Rat et al., 2014). Four trials were conducted with melanoma patients (Bowen et al., 2015; Robinson et al., 2010; Robinson et al., 2007a; Turrisi et al., 2015), 2 trials with first degree relatives of melanoma patients (Geller et al., 2006; Manne et al., 2010), 6 trials with at-risk outpatients (Glanz et al., 2010; Glanz et al., 2015; Glanz et al., 2013; Glazebrook et al., 2006; Oliveria et al., 2004; Rat et al., 2014), and 1 trial with at-risk young adults (Heckman et al., 2016). Only 6 of the 13 trials were registered with trial registry platforms (Glanz et al., 2015; Heckman et al., 2016; Manne et al., 2010; Rat et al., 2014; Robinson et al., 2007a; Turrisi et al., 2015).

3.3. Outcome 1: SSE Conceptualization

Consistent with our inclusion criteria, all 13 trials provided recommendations for SSE. However, sufficient detail about the SSE recommendations in terms of conceptualization of behavior were reported in only 6 trials (Glanz et al., 2015; Heckman et al., 2016; Oliveria et al., 2004; Robinson et al., 2007a; Robinson et al., 2010; Turrisi et al., 2015). In all 6 trials, SSE denoted checking the skin on the entire body for the early signs of skin cancer using the ABCDE criteria. These trials also recommended the use of tools to visualize (magnifying lens, mirrors), measure (ruler), and/or track changes in the lesions (body map, score card, whole body photographs). The recommended frequency for SSE was reported as “monthly” in 3 trials (Glanz et al., 2015; Robinson et al., 2010; Turrisi et al., 2015). The remainder of 10 trials did not report the recommended frequency for SSE provided to participants.

3.4. Outcome 2: SSE Measurement and Consistency of SSE Methods

The number of items used to assess SSE behavior ranged from 1 item (Glanz et al., 2010; Heckman et al., 2016; Manne et al., 2010; Oliveria et al., 2004) to 17 items (Turrisi et al., 2015). The items reflected individually-performed self-exams (or SSE) (Bowen et al., 2015; Geller et al., 2006; Glanz et al., 2010; Glanz et al., 2015; Glanz et al., 2013; Glazebrook et al., 2006; Manne et al., 2010; Oliveria et al., 2004; Rat et al., 2014; Robinson et al., 2010; Robinson et al., 2007a), partner-assisted skin exams (Geller et al., 2006; Glazebrook et al., 2006; Rat et al., 2014; Robinson et al., 2010; Robinson et al., 2007a; Turrisi et al., 2015), and skin exams conducted with tracking and/or monitoring devices such as body maps (or SSE with tracking/monitoring) (Geller et al., 2006; Rat et al., 2014; Robinson et al., 2010; Robinson et al., 2007a).

At the content level, based on how the items were phrased, we identified 3 separate modalities to assess skin-self exams: a) as an exam performed or not (Bowen et al., 2015; Geller et al., 2006; Glanz et al., 2010; Glanz et al., 2013; Glazebrook et al., 2006; Heckman et al., 2016; Oliveria et al., 2004; Rat et al., 2014); b) as frequency or recency of the (last) exam (Glanz et al., 2015; Glanz et al., 2013; Manne et al., 2010; Robinson et al., 2007a; Robinson et al., 2010); and c) as comprehensiveness of the exam (extent or number of body parts examined) (Bowen et al., 2015; Turrisi et al., 2015).

The SSE behavior outcomes were scored as proportion of individuals endorsing the behavior (binary variable) (Bowen et al., 2015; Geller et al., 2006; Glanz et al., 2010; Glazebrook et al., 2006; Heckman et al., 2016; Oliveria et al., 2004; Rat et al., 2014) or as mean and standard deviation of a specific range of behavior (continuous variable) (Glanz et
al., 2015; Glanz et al., 2013; Manne et al., 2010; Robinson et al., 2010; Robinson et al., 2007a; Turrisi et al., 2015). For the binary outcomes, 6 trials conceptualized the self-exam as “at least 1 SSE during [specified timeframe]” (Bowen et al., 2015; Geller et al., 2006; Glanz et al., 2010; Glazebrook et al., 2006; Heckman et al., 2016; Rat et al., 2014) whereas one trial used a different definition “at least 3 SSE’s in the last 4 months” (Oliveria et al., 2004). For the continuously scored outcomes, the answer choices indicated different frequency ranges for SSE across different trials (specific details included in Table 2). The timeframe for the assessment of SSE behavior ranged from “last 2 months” (Bowen et al., 2015) to “last 12 months” (Geller et al., 2006; Rat et al., 2014), with the most frequent timeframe being “last 4 months” (Oliveria et al., 2004; Robinson et al., 2010; Robinson et al., 2007a; Turrisi et al., 2015).

Generally, there was substantial variability between trials with respect to the item content, response format, scoring procedures, and timeframe of assessment for SSE measures, with no two trials using identical methods. Notably, there were more similarities in item content among trials that used binary variables and more differences among trials that used continuous variables.

3.5. Outcome 3: Validity and Reliability of SSE Measures

All of the SSE measures (or items) used in the trials included in this review (n = 13) showed face validity, as the items were directly relevant to the assessment of self-checking the skin for signs of skin cancer and had been previously used in melanoma prevention research. We could not identify any evidence supporting the content (beyond face validity), construct, or criterion validity of the SSE instruments used in the trials of this review. With respect to reliability, only 1 of the 13 trials reported internal consistency reliability coefficients for the SSE behavior items (Turrisi et al., 2015). The exact coefficients were included in Table 2.

4. Discussion

We reviewed aspects of assessing SSE from 13 unique behavioral trials. With respect to our first research question, how was SSE defined, which was extrapolated from the instructions provided to study participants, we found that fewer than half of the eligible trials reported the specific instructions about SSE. Among the trials that reported such information, SSE behavior involved checking the entire body for the early signs of skin cancer, including the areas located on the back, with the help from a partner or by using two mirrors, while also using aids to visualize, measure, and track moles. The recommended criteria to use for identifying problematic skin lesions was the ABCDE criteria (Abbasi et al., 2004). We found no references to the other two criteria typically used for melanoma early detection, the ugly duckling (Grob & Bonerandi, 1998) or the Glasgow 7-point checklist (MacKie et al., 2007). The recommended frequency for SSE behavior, which was reported in only 3 of 13 trials, was “monthly SSE”. At a minimum, an intervention designed to promote SSE as a melanoma prevention behavior should necessarily include several key aspects, as illustrated by the Information-Motivation-Behavioral Model (Fisher & Fisher, 1992; 2002) which is an atheoretical, pragmatic model for predicting health behavior change: a)
identify the preventive health behavior as an important aspect of prevention (information) and educate the target population about the benefits of performing such a behavior (psychoeducation), b) motivate individuals to engage in the behavior, by means of creating acceptable personal, social, and professional norms/attitudes around the behavior (motivation); and c) illustrate how to perform the behaviour and the recommended frequency (behavioral skills), in the case of SSE: how to check the skin for the early signs of melanoma and how often (e.g., monthly, yearly). Every behavioral component of an intervention is important to state clearly, so that participants know what is expected of them and researchers can assess whether there was consistency between the behaviors recommended to participants and the assessment of these behaviors post intervention. Finally, complete and transparent reporting of interventions, including all active components or ingredients, as per CONSORT-SPI (Montgomery et al., 2018) or TIDieR (Hoffmann et al., 2014) guidelines of reporting, reduces waste in research, by providing adequate information for applied clinicians and for researchers involved in synthesizing or meta-analyzing the effect of interventions and in designing replication studies (Glasziou et al., 2014).

With respect to our second research question, how was SSE measured in behavioral trials with individual at elevated risk for melanoma and the consistency of assessment methods across trials, we found substantial variability at the level of item content, specifically with respect to what the items inquired about, i.e., whether a skin self-exam was performed or not, what was the frequency or recency of the last skin exam performed, or how many body parts were examined. Also, there were inconsistencies related to the number of items used to assess SSE, with an increased number of items being indicative of a larger extent of skin being covered by the exam and/or the complexity of the behavior assessed (e.g., partner assistance, use of tracking or monitoring devices). We found variability in response format (scales and anchors), a wide range covering the time frame of the assessment (2 months to 12 months), and two types of variables used in analyses, proportions and means/standard deviations. The lack of conceptual consistency with respect to SSE measurement is not surprising given that currently there are no empirical studies (randomized or using other designs) that investigated whether different modalities of self-examining the skin led to differential health (e.g., mortality) or health-related (e.g., early detection) outcomes. In addition, the inconsistencies identified in this review may reflect inconsistencies found across the clinical practice guidelines for melanoma prevention and follow-up care. Specifically, guideline recommendations for SSE for high-risk groups range from physicians’ responsibility to raise awareness about the importance to SSE to physicians (or nurses) giving live demonstrations on how to check the skin for problematic lesions accompanied by written materials that patients can consult at home whole conducting SSE's (Marciano et al., 2014; Watts et al., 2015). Guideline recommendations about how frequently SSE should be performed are also highly inconsistent ranging from monthly (Bichakjian et al., 2011; Marsden et al., 2010) to every 3-6 months (Australian Cancer Network Melanoma Guidelines Revision Working Party, 2008; Institut National du Cancer (INCa) and Haute Autorite de Sante (HAS), 2012) or are simply not specified (National Collaborating Centre for Cancer, 2015).
Future research will benefit from establishing the most efficacious modality to perform SSE as well as how often SSE should be performed in order to effectively identify problematic lesions.

Finally, our third research question investigated the level of evidence supporting the validity and reliability of the instruments assessing SSE behavior. We found no evidence supporting the validity of SSE measures used across trials, beyond face validity. We found support for the internal consistency reliability of an SSE measure used in one trial. In a few instances, we found articles using measures similar or identical to the ones used in some of the trials included in this review (Bowen et al., 2015; Manne et al., 2010; Oliveria et al., 2004). These additional studies investigated base rates, correlates, and predictors of SSE (Manne et al., 2004; Robinson, Fisher, & Turrisi, 2002; Weinstock et al., 1999; Weinstock et al., 2004), but did not directly address the measurement properties of any of the SSE instruments used.

As reported, there is variability in the SSE intervention literature with respect to the conceptualization of SSE and there are inconsistencies across studies in how SSE is measured. It might seem intuitive that a standard SSE should involve a deliberate/planned whole-body skin exam plus the use of external aids, such as a partner’s help or two mirrors to check inaccessible areas, and that such an exam could be labelled as a thorough or complete SSE (Weinstock et al., 2004) as opposed to casual, once in a while, opportunistic checking of the skin. One might further assume that “thorough SSE” should be recommended to all of the individuals at-risk for melanoma. Such a definition could in theory be adopted by the melanoma research community. However, the reality is that the current melanoma prevention literature does not provide sufficient evidence to conclude that one modality of conceptualizing and measuring SSE is more effective than others at predicting health and health-related outcomes. We need to learn more about SSE, including 1) what is the most pertinent definition of SSE behavior? 2) is there a valid and reliable method to assess SSE behavior? and 3) are some modalities of assessing SSE more effective than others at predicting health outcomes in individuals at increased risk for melanoma? Answers to these questions have important implications for several reasons. First, the way SSE is defined affects the reported SSE rates in the literature. For instance, when SSE behavior was defined as any skin checking (as opposed to none), SSE rates ranged from 46% among the general population (Robinson, Rigel, & Amonette, 1998) to 70% among a population at increased risk for melanoma (Robinson et al., 2002), and between 75-90% among individuals with a personal history of melanoma (Kasparian et al., 2012). When SSE behavior was defined as a deliberate and systematic behavior of checking the skin, reported rates ranged from 12% throughout one’s lifetime to 18% during the past two months among an outpatient population without an increased risk for melanoma (Weinstock et al., 2004). Second, the inconsistent definition and measurement of SSE in tandem with a lack of evidence supporting the validity and reliability of existing SSE measures affect one’s confidence in the interpretation, comparison, and synthesis of the results from randomized controlled trials using SSE as an outcome.
4.1. Limitations

First, we used strict population inclusion (i.e., with an increased risk of melanoma compared to the general population) and exclusion (general population and white populations without an elevated risk for melanoma) criteria, as guided by the primary research question of this systematic review, which was to investigate the effect of behavioral interventions on mortality, melanoma early detection, and SSE uptake in high-risk populations. Our population inclusion criteria align with most melanoma care guidelines, which recommend SSE only for individuals with an increased risk of melanoma incidence compared to the general population (for reviews of guidelines for at-risk individuals, see (Marciano et al., 2014; Watts et al., 2015). Nonetheless, these methodological decisions limit the generalizability of our results to the populations we excluded. Second, we only included RCT’s and excluded studies with all other research designs, with the aim of identifying preventative interventions with the highest level of evidence support. In doing so, it is possible we missed relevant interventions in the early stage of development and testing. Last, we acknowledge that our search strategy was not built specifically for the purpose of identifying SSE measures and their psychometric properties. Nevertheless, given the preparation, rigour, and resources involved in conducting clinical trials, it is reasonable to expect that if valid and reliable SSE measures existed, they would have been used in at least some of these trials and adequate references would have been included in the trial reports. Hence, it seems reasonable to assume that we reviewed a representative sample of SSE measures that have undergone a more stringent measure selection, as they served as outcome variables in resource-intensive randomized controlled trials.

4.2. Future Directions for Research and Practice Implications

Future research should first establish the construct validity of scales measuring SSE behavior. One possible approach might be a large Delphi study with melanoma prevention experts, researchers and methodologists, and patient representatives to reach consensus on the most relevant items to be included in such a measure. As a second step, it would be important for future empirical studies to investigate the predictive validity of SSE measures. Specifically, determining whether deliberate checking of the skin of the entire body (or of moles pointed out by health care professionals as needing surveillance) differs from casually checking the skin, and whether the use of tracking devices adds a substantial benefit with respect to short and long-term outcomes, such as SSE uptake and melanoma early detection. Such research projects could answer questions about the optimal frequency of SSE behavior, which could then be incorporated into guidelines of care for individuals at increased risk for melanoma, which is much needed given that the current guidelines recommendations for SSE are mostly consensus-based (Marciano et al., 2014). Ultimately, agreement on SSE definition and most adequate measurement method among the melanoma scholarly community could lay the groundwork for studies testing the impact of SSE behavior on the early detection of melanoma and melanoma-related mortality.
4.3. Conclusion

The current review of behavioral interventions for individuals at increased risk for melanoma found that the outcome measure, skin self-examination (SSE), is operationalized (defined and measured) inconsistently across different randomized controlled trials. Further, there is minimal evidence supporting the validity and reliability of the SSE measures used across these trials. Future research is needed to investigate the psychometric properties of SSE measures, including establishing the predictive validity of such scales by examining the impact of SSE on health outcomes in populations at-risk for melanoma.

Word count: 5,178 words (excluding title page, abstract, acknowledgment, references, tables, appendices).
Characters: 28,990 (excluding title page, abstract, acknowledgment, references, tables, appendices).
Acknowledgment

We would like to thank Jill Boruff, MLIS, Associate Librarian at McGill Library and Archives, who peer-reviewed the Medline search, which we used as a basis to build the searches in all of the other databases.
References


Table 1. Characteristics of Trials of Behavioral Interventions with Individuals at Increased Risk for Melanoma

<table>
<thead>
<tr>
<th>First author, year</th>
<th>Country</th>
<th>Funding source</th>
<th>Trial registration, acronym</th>
<th>Population</th>
<th>Intervention vs. control</th>
<th>N Randomized</th>
<th>Age M ± SD or n (%)</th>
<th>Female n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bowen, 2015, 2019</td>
<td>United States (Bowen et al., 2015; 2019)</td>
<td>NIH# P20 HG007343 NCI# R01 CA107430</td>
<td>Unregistered trial Suntalk Study</td>
<td>Melanoma patients FDR of melanoma patientsb</td>
<td>Web-based communication and support vs. Wait-list</td>
<td>Intx 157ab Ctrl 156ab</td>
<td>56.11 ± 12.33a 51.32 (Not reportedb)</td>
<td>Total 173 (56)a Total 199 (63.6)b</td>
</tr>
<tr>
<td>2. Geller, 2006, United States (Geller et al., 2006)</td>
<td>NIH# R01CA76333</td>
<td>Unregistered trial</td>
<td></td>
<td>FDR of melanoma patients</td>
<td>Tailored phone counselling vs. TAU</td>
<td>Intx 237 Ctrl 257</td>
<td>Reported &gt; 51 Intx 105 (44.3) Ctrl 101 (39.4)</td>
<td>Intx 123 (51.9) Ctrl 141 (54.9)</td>
</tr>
<tr>
<td>3. Glanz, 2010, United States (Glanz et al., 2010)</td>
<td>NCICA 76419</td>
<td>Unregistered trial Project SCAPE</td>
<td></td>
<td>At-risk outpatients (identified with BRAT)</td>
<td>Tailored intervention vs. TAU</td>
<td>Intx 362 Ctrl 362</td>
<td>Intx 42.1 ± 10.8 Ctrl 41.2 ± 11.2</td>
<td>Intx 285 (78.7) Ctrl 276 (76.2)</td>
</tr>
<tr>
<td>4. Glanz, 2013, United States (Glanz et al., 2013)</td>
<td>NIH# 5UC2CA148310-02</td>
<td>Unregistered trial GenoMEL Cohort</td>
<td></td>
<td>At-risk outpatients (melanoma patients or FDR)</td>
<td>Genetic counselling vs. TAU</td>
<td>Intx 35 Ctrl 38</td>
<td>Intx 62.4 ± 14.6 Ctrl 56.9 ± 16.1</td>
<td>Intx 28 (80.0) Ctrl 22 (57.9)</td>
</tr>
<tr>
<td>5. Glanz, 2015, United States (Glanz et al., 2015)</td>
<td>NIH# 5UC2 CA148310 NCT01356771</td>
<td>PennSCAPE trial</td>
<td></td>
<td>At-risk outpatients (identified with BRAT)</td>
<td>Tailored mailing vs. Generic mailing</td>
<td>Intx 95 Ctrl 111</td>
<td>Intx 53.5 ± 14.4 Ctrl 56.5 ± 15.7</td>
<td>Intx 62 (74.7) Ctrl 79 (72.5)</td>
</tr>
<tr>
<td>6. Glazebrook, 2006, United Kingdom (Glazebrook et al., 2006)</td>
<td>Trent NHS</td>
<td>Unregistered trial Skinsafe trial</td>
<td></td>
<td>At-risk outpatients (family or personal history, phenotypic features, skin type, and number of nevi)</td>
<td>Health education via computer program vs. TAU</td>
<td>Intx 259 Ctrl 330</td>
<td>Intx 38.2 ± 14.3 Ctrl 38.4 ± 15.2</td>
<td>Intx 214 (82.6) Ctrl 259 (78.5)</td>
</tr>
<tr>
<td>7. Heckman, 2016, United States (Hedman et al., 2016)</td>
<td>NIH# 1R01CA154928-01 NCT02147080 UV4me</td>
<td></td>
<td></td>
<td>At-risk young adults (identified with BRAT)</td>
<td>Tailored website vs. Generic website vs. Assessment only</td>
<td>Intx 287 Ctrl1 338 Ctrl2 337</td>
<td>Intx 21.8 ± 2.2 Ctrl1 21.7 ± 2.2 Ctrl2 21.9 ± 2.2</td>
<td>Intx 197 (68.6) Ctrl1 216 (64.1) Ctrl2 224 (65.9)</td>
</tr>
<tr>
<td>8. Manne, 2010, United States (Manne et al., 2010)</td>
<td>No funding reported NCT00816374</td>
<td></td>
<td></td>
<td>FDR of melanoma patients</td>
<td>Tailored print and counselling vs. Generic print and counselling</td>
<td>Intx 225 Ctrl 218</td>
<td>Intx 48.1 ± 12.6 Ctrl 47.1 ± 13.9</td>
<td>Intx 135 (60.0) Ctrl 144 (66.1)</td>
</tr>
<tr>
<td>9. Oliveria, 2004, United States (Oliveria et al., 2004)</td>
<td>No funding reported Unregistered trial</td>
<td></td>
<td></td>
<td>At-risk outpatients (&gt; = 5 clinical dysplastic or atypical nevi)</td>
<td>Teaching with photo-book vs. Teaching without photo-book</td>
<td>Intx 49 Ctrl 51</td>
<td>Intx 40.3 ± 10.9 Ctrl 39.4 ± 11.5</td>
<td>Intx 29 (59.2) Ctrl 34 (66.7)</td>
</tr>
<tr>
<td>First author, year</td>
<td>Country</td>
<td>Funding source</td>
<td>Trial registration, acronym</td>
<td>Population</td>
<td>Intervention vs. control</td>
<td>N Randomized</td>
<td>Age M ± SD or n (%)</td>
<td>Female n (%)</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>----------------</td>
<td>----------------------------</td>
<td>------------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>---------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>10. Rat, 2014, France (Rat et al., 2014)</td>
<td>10.. Rat, 2014, France (Rat et al., 2014)</td>
<td>No funding reported</td>
<td>NCT01610531 “COPARIME project”</td>
<td>At-risk outpatients (identified with SAMScore)</td>
<td>Targeted screening and education vs. TAU</td>
<td>Intx 97 Ctrl 76</td>
<td>Intx 43.6 ± 17.1 Ctrl 42.8 ± 14.6</td>
<td>Intx: 74 (76.0) Ctrl: 58 (76.0)</td>
</tr>
<tr>
<td>11. Robinson, 2007, United States (Robinson et al., 2007a)</td>
<td>11.. Robinson, 2007, United States (Robinson et al., 2007a)</td>
<td>NCI# 5R21 CA-103833-02 NCT01013844</td>
<td>Melanoma patients</td>
<td>Dyadic learning vs. Solo-learning</td>
<td>Intx 65 Ctrl 65</td>
<td>(Median) Intx 40-49 Ctrl 50-59</td>
<td>Intx: 33 (51.0) Ctrl: 32 (49.0)</td>
<td></td>
</tr>
<tr>
<td>12. Robinson, 2010, United States (Robinson et al., 2010)</td>
<td>12.. Robinson, 2010, United States (Robinson et al., 2010)</td>
<td>No funding reported</td>
<td>Unregistered, but declared as Pilot for Turrisi trial</td>
<td>Melanoma patients</td>
<td>Intervention via workbook vs. In-person intervention</td>
<td>Intx 21 Ctrl 19</td>
<td>(Median) Intx 40-49 Control 40-49</td>
<td>Intx: 10 (53.0) Ctrl: 11 (52.0)</td>
</tr>
<tr>
<td>13. Turrisi, 2015, United States (Turrisi et al., 2015)</td>
<td>13.. Turrisi, 2015, United States (Turrisi et al., 2015)</td>
<td>NCI# R01 CA134908 NCT01432860</td>
<td>Melanoma patients</td>
<td>In-person intervention vs. Intervention via workbook vs. Intervention via e-Tablet vs. TAU</td>
<td>In-person 165 Workbook 159 Tablet 71</td>
<td>(Median) In-person 50-59 Workbook 50-59 Tablet 50-99</td>
<td>In person 75 (45.5) Workbook 81 (50.9) Tablet 37 (52.1) Table 60 (60.6)</td>
<td></td>
</tr>
</tbody>
</table>

Note. NIH = National Institutes of Health; NCI = National Cancer Institute; FDR = first degree relative; SSE = skin self-examination; TAU = treatment-as-usual; BRAT=Brief skin cancer risk assessment tool; SAMScore = the Self-Assessment Melanoma Risk Score; Intx = intervention; Ctrl = control

a Results pertaining to melanoma patients
b Results pertaining to FDR of melanoma patients
<table>
<thead>
<tr>
<th>First author, year</th>
<th>SSE Recommendations</th>
<th>Frequency</th>
<th>Items</th>
<th>Response options</th>
<th>Time frame</th>
<th>Scoring</th>
<th>SSE</th>
<th>Outcome definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bowen, 2013, 2019 United States (Bowen et al., 2015; 2019)</td>
<td>Details not reported.</td>
<td>Not reported</td>
<td>1-7. Did you carefully examine … [7 body areas] = (SSE - comprehensive)</td>
<td>Never</td>
<td>Last 2 months</td>
<td>Binary (%)</td>
<td>SSE</td>
<td>At least 1x SSE (body part) in the last 2 months</td>
</tr>
<tr>
<td>2. Geller, 2006 (Geller et al., 2006)</td>
<td>Details not reported.</td>
<td>Not reported</td>
<td>8. Did you do a thorough skin examination? (SSE)</td>
<td>Yes</td>
<td></td>
<td></td>
<td>SSE</td>
<td>At least 1 thorough SSE in the last 2 months</td>
</tr>
<tr>
<td>3. Glanz, 2010 (Glanz et al., 2010)</td>
<td>Details not reported.</td>
<td>Not reported</td>
<td>1. Have you ever/ in the last 3 months conducted a thorough SSE? (SSE)</td>
<td>Yes</td>
<td>Last 3 months</td>
<td>Binary (%)</td>
<td>SSE</td>
<td>At least 1x SSE in the last 3 months</td>
</tr>
<tr>
<td>4. Glanz, 2013 (Glanz et al., 2013)</td>
<td>Details not reported.</td>
<td>Not reported</td>
<td>1. Do you ever closely examine yourself for signs of skin cancer, including melanoma? (SSE)</td>
<td>Yes</td>
<td>Ever</td>
<td>Binary (%)</td>
<td>SSE</td>
<td>Recency of last thorough SSE (“never” to “last month”)</td>
</tr>
<tr>
<td>5. Glanz, 2015 (Glanz et al., 2015)</td>
<td>Instructions on how to perform SSE using ABCDE criteria; how to use mirrors for back areas and scalp; how to track changes in moles included in intx and ctrl.</td>
<td>Monthly SSE</td>
<td>1. Have you ever closely examined yourself skin for signs of skin cancer, including melanoma? (SSE)</td>
<td>Items combined: 1 = Never 2 = &gt; 3 months ago</td>
<td>Continuous (M/SD)</td>
<td></td>
<td>SSE</td>
<td>Recency of last SSE (“never” to “last month”)</td>
</tr>
<tr>
<td>First author, year</td>
<td>SSE Recommendations</td>
<td>Frequency</td>
<td>SSE Assessment</td>
<td>Outcome definition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------</td>
<td>-----------</td>
<td>---------------</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Glazebrook, 2006 (Glazebrook et al., 2006)</td>
<td>Details not reported.</td>
<td>Not reported</td>
<td>1. [Have participants been] “checking moles”? (SSE)</td>
<td>SSE At least 1 x SSE in the last 6 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. [Have the participants performed] SSE? (SSE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. [Have the participants] had their skin checked by someone else? (PASE)</td>
<td>PASE At least 1 x PASE in the last 6 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Heckman, 2016 (Heckman et al., 2016)</td>
<td>Instructions on how to perform SSE using the ABCDE criteria + barriers to SSE.</td>
<td>Not reported</td>
<td>1. [In the last 3 months/ since last assessment] have you or your partner examined your entire body, including your back, for skin cancer? (SSE/PASE)</td>
<td>SSE At least 1 x SSE in the last 3 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Manne, 2010 (Manne et al., 2010)</td>
<td>Details not reported.</td>
<td>Not reported</td>
<td>1. How often have [participants] examined their skin deliberately and purposefully in the past year/ last 3 months? (SSE)</td>
<td>SSE Frequency of SSE in the last 3 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Oliveria, 2004 (Oliveria et al., 2004)</td>
<td>Generic video on SSE how-to (systematic checking of skin on the entire body) and nurse-led demo on SSE (systematic checking of entire body using ABCDE criteria to identify problematic lesions) included in intx and ctrl. Individual, whole-body photographs (to track changes) included in intx.</td>
<td>Not reported</td>
<td>1. How many times in the past 4 months did you (or someone else) usually, thoroughly examine your skin? By thorough we mean actually looking at different areas of your skin deliberately and systematically. (SSE)</td>
<td>SSE At least 3 x SSE in the last 4 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Rat (Rat et al., 2014)</td>
<td>Details not reported.</td>
<td>Not reported</td>
<td>1. In the last 12 months, did you perform a skin self-examination? (SSE)</td>
<td>SSE At least 1 x SSE in the last 12 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. If yes, did you have assistance from another person or mirror? (PASE)</td>
<td>PASE At least 1 x PASE in the last 12 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First author, year</td>
<td>SSE Recommendations</td>
<td>SSE Assessment</td>
<td>SSE Outcome definition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------</td>
<td>----------------</td>
<td>------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| **11. Robinson, 2007** (Robinson et al., 2007a) | Card with ABCDE criteria and SSE RA-led demo of SSE how-to (e.g., using magnifying glass for criterion E) included in intx and ctrl. | 3. If yes, did you take a photograph? (SSE with tracking/monitoring) | SSE w/ tracking  
At least 1x SSE w/ photographs in the last 12 months |
| | Not reported | 1. [How many] times patients examined their skin? (SSE) | SSE | SSE w/ tracking  
At least 1x SSE w/ photographs in the last 12 months |
| | | 2. [How many] times patients examined their skin with a partner? (PASE) | PASE | Frequency of PASE in the last 4 months (*never* to *daily*) |
| | | 3. [...use of the] Body Map? (SSE with tracking/monitoring) | SSE | Frequency of SSE in the last 4 months (*never* to *daily*) |
| **12. Robinson, 2010** (Robinson et al., 2010) | Self-screening kit (ruler, ABCDE card, magnifying glass, body map) and instructions on how to use it included in intx and ctrl. | Monthly SSE | PASE | Frequency of PASE in the last 4 months (*never* to *daily*) |
| | Monthly SSE | 1. How often does the participant perform SSE with a partner? (PASE) | SSE | Frequency of SSE in the last 4 months (*never* to *daily*) |
| **13. Turrisi, 2015** (Turrisi et al., 2015) | ABCDE criteria explained and self-screening kit (lighted magnifying lens, laminated ABCDE card, body map, and scorecards to track monthly SSE) included in intx and ctrl. | Monthly SSE | PASE | Frequency of PASE in the last 4 months (*never* to *daily*) |

Note. SSE = skin self-examination; PASE = partner-assisted SSE; Intx = intervention; Ctrl = control.

* Trained RA presented the ABCDE info on the card; showed 2 additional color printed examples of E (evolution) of pigmented lesions in 1 year; answered questions about the card content; gave a skills demonstration using a magnifying lens to look at moles; and pointed out irregular borders and uneven colors.

+ Bowen, 2015 (Bowen et al., 2015): 7 items corresponding to 7 body areas (the front of you from waist up, the front of your thighs and legs, the bottom of your feet, your calves, the back of thighs and legs, the buttocks and lower part of your back, your upper back).

+ Total body SSE Cronbach’s alpha = .96-.98; Total body SSE = 17 items corresponding to 17 different body areas (e.g., front and back of neck; top and soles of feet; front and back of thighs; groin).

+ SSE of easy to see areas Cronbach’s alpha = 0.95-0.98; SSE of easy to see areas = 7 items corresponding to 7 body areas (face, front torso, neck, both hands, arms, legs, and feet).

+ SSE of difficult to see areas Cronbach’s alpha = 0.88-0.95; SSE of difficult to see areas = 6 items corresponding to 7 body areas (scalp, buttocks, back of ears, neck, shoulders, thighs).
Figure 1. PRISMA flowchart of study selection.

*Total records included in this review N = 21: N = 15 publications, pertaining to 13 unique trials; N=6 trial registrations

**Details of the manual searches and the reasons for exclusion are included in Appendix D.
## Appendix A: Medline Search

Medline

Database(s): Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily, Ovid MEDLINE(R) and Ovid OLDMEDLINE(R) 1946 to Present

Search Strategy:

<table>
<thead>
<tr>
<th>#</th>
<th>Searches</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Melanoma/ or melanoma.mp.</td>
<td>102432</td>
</tr>
<tr>
<td>2</td>
<td>self examination.mp. or Self-Examination/</td>
<td>3558</td>
</tr>
<tr>
<td>3</td>
<td>self exam*.mp.</td>
<td>3707</td>
</tr>
<tr>
<td>4</td>
<td>self screen*.mp.</td>
<td>151</td>
</tr>
<tr>
<td>5</td>
<td>self surveillance.mp.</td>
<td>57</td>
</tr>
<tr>
<td>6</td>
<td>self monitor*.mp.</td>
<td>8813</td>
</tr>
<tr>
<td>7</td>
<td>skin exam*.mp.</td>
<td>769</td>
</tr>
<tr>
<td>8</td>
<td>2 or 3 or 4 or 5 or 6 or 7</td>
<td>13364</td>
</tr>
<tr>
<td>9</td>
<td>skin neoplasms.mp. or Skin Neoplasms/</td>
<td>101285</td>
</tr>
<tr>
<td>10</td>
<td>skin cancer.mp.</td>
<td>13147</td>
</tr>
<tr>
<td>11</td>
<td>skin surveillance.mp.</td>
<td>41</td>
</tr>
<tr>
<td>12</td>
<td>8 or 11</td>
<td>13396</td>
</tr>
<tr>
<td>13</td>
<td>1 or 9 or 10</td>
<td>174715</td>
</tr>
<tr>
<td>14</td>
<td>12 and 13</td>
<td>766</td>
</tr>
<tr>
<td>15</td>
<td>randomized controlled trial.pt.</td>
<td>404274</td>
</tr>
<tr>
<td>16</td>
<td>controlled clinical trial.pt.</td>
<td>89994</td>
</tr>
<tr>
<td>17</td>
<td>randomi?ed.ab.</td>
<td>397618</td>
</tr>
<tr>
<td>18</td>
<td>placebo.ab.</td>
<td>165263</td>
</tr>
<tr>
<td>19</td>
<td>drug therapy.fs.</td>
<td>1810175</td>
</tr>
<tr>
<td>20</td>
<td>randomly.ab.</td>
<td>240295</td>
</tr>
<tr>
<td>21</td>
<td>trial.ab.</td>
<td>343550</td>
</tr>
<tr>
<td>22</td>
<td>groups.ab.</td>
<td>1504295</td>
</tr>
<tr>
<td>23</td>
<td>or/15-22</td>
<td>3643127</td>
</tr>
<tr>
<td>24</td>
<td>exp animals/ not humans.sh.</td>
<td>4173179</td>
</tr>
<tr>
<td>25</td>
<td>23 not 24</td>
<td>3133238</td>
</tr>
<tr>
<td>26</td>
<td>14 and 25</td>
<td>187</td>
</tr>
</tbody>
</table>
## Appendix B: Inclusion and Exclusion Criteria for Relevant Trials

<table>
<thead>
<tr>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(P) Population:</strong></td>
<td><strong>Exclusion:</strong></td>
</tr>
<tr>
<td>- Individuals at increased risk for melanoma</td>
<td>- General Population</td>
</tr>
<tr>
<td>- Adults (18+)</td>
<td>- Children/youth below 18</td>
</tr>
<tr>
<td><strong>Increased Risk (environmental + bio factors):</strong></td>
<td>- Men 50+</td>
</tr>
<tr>
<td>- Immunosuppressed individuals (radiation; transplant)</td>
<td>- Spouses and other non-blood relatives of melanoma patients</td>
</tr>
<tr>
<td>- History of childhood radiation</td>
<td><strong>Other examples of non-eligible groups:</strong></td>
</tr>
<tr>
<td>- Melanoma patients</td>
<td>- Beach-goers</td>
</tr>
<tr>
<td>- People w/ a history of melanoma</td>
<td>- University/college students</td>
</tr>
<tr>
<td>- Keratinocyte cancer (NMSC: BCC/SCC) patients</td>
<td>- Patients presenting for routine medical visits</td>
</tr>
<tr>
<td>- 100+ naevi (moles; birth or beauty marks) or 5+ atypical naevi</td>
<td>- Healthcare professionals working in melanoma clinics</td>
</tr>
<tr>
<td>- First degree relatives of melanoma patients</td>
<td>- Parents of children treated for childhood cancers</td>
</tr>
<tr>
<td>- 50-100 naevi; 1+ atypical naevi</td>
<td><strong>(I) Intervention:</strong></td>
</tr>
<tr>
<td>- Phenotypic features</td>
<td>- Design: Not an RCT</td>
</tr>
<tr>
<td>- red or blond hair</td>
<td>- Content: Medical procedures</td>
</tr>
<tr>
<td>- tendency to freckle</td>
<td>(pharmacological, surgery/biopsy)</td>
</tr>
<tr>
<td>- gets easily burnt</td>
<td>- <strong>(C) Comparison:</strong></td>
</tr>
<tr>
<td>- tanned skin</td>
<td>- at least one other intervention or control arm</td>
</tr>
<tr>
<td><strong>(O) Outcomes:</strong></td>
<td><strong>(O) Outcomes:</strong></td>
</tr>
<tr>
<td>- SSE behavior</td>
<td>- SSE not an outcome</td>
</tr>
<tr>
<td>- Early detection</td>
<td>- SSE Accuracy as outcome (i.e., identifying problematic lesions accurately via SSE)</td>
</tr>
<tr>
<td>- Mortality</td>
<td>- Any other outcomes</td>
</tr>
<tr>
<td><strong>Language:</strong></td>
<td><strong>Language:</strong></td>
</tr>
<tr>
<td>- English, French, German, Spanish, Russian, Romanian</td>
<td>- Any other language</td>
</tr>
</tbody>
</table>

*Note. NMSC = non-melanoma skin cancer; BCC = basal cell carcinoma; SCC = squamous cell melanoma*
Appendix C: Data Coding Sheet for Full Text Review

Ref: 1. There and Back Again: A Review of Residency and Return Migrations in Sharks, with Implications for Population Structure and Management. Chapman DD, Feldheim KA, Papastamatiou Y, Hueter RE.

The overexploitation of sharks has become a global environmental issue in need of a comprehensive and multifaceted management response. Tracking studies are beginning to elucidate how shark movements shape the internal dynamics and structure of populations, which determine the most appropriate scale of these management efforts.

Tracked sharks frequently either remain in a restricted geographic area for an extended period of time or residency prior to a previously resided-in area after making long-distance movements (site fidelity). Genetic studies have shown that some individuals of certain species preferentially return to their natal birthplace (natal philopatry) or birth region (regional philopatry) for either parturition or mating, even though they make long-distance movements that would allow them to breed elsewhere. More than 80 peer-reviewed articles, constituting the majority of published shark tracking and population genetic studies, provide evidence of at least one of these behaviors in a combined 31 shark species from six of the eight extant orders.

Residency, site fidelity, and philopatry can alone or in combination structure many coastal shark populations on finer geographic scales than expected based on their potential for dispersal. This information should therefore be used to scale and inform assessment, management, and conservation activities intended to restore depleted shark populations. Expected final online publication date for the Annual Review of Marine Science Volume 7 is January 03, 2015.

1. Population Adults at Increased Risk for Melanoma?
   - Yes
   - No

2. Intervention type RCT
   - Intervention contents: psycho-educational, informational
     - Yes
     - No

3. Outcome SSE
   - Yes
   - No

4. Outcome: Early Detection?
   - (e.g., melanoma size, depth; melanoma stage)
     - Yes
     - No

5. Outcome: Mortality or Survival?
   - Yes
   - No

6. Entry in Trial Registry?
   - Yes
   - No

7. Relevant Trial Registry Entry?
   - Yes
   - No
   - N/A
Appendix D: Results of Manual Searches of Reference Lists of Included Trial Reports

Summary of manual searches: N=25 articles referenced in the method sections/measures of the included trials:
N=3 duplicate sources
N=4 already included trial reports
N=6 not about SSE
N=12 correlates or predictors of SSE, but not study aim to establish validity of SSE measures

<table>
<thead>
<tr>
<th>Included trial</th>
<th>Sources cited in method section of included trial report</th>
</tr>
</thead>
</table>
| Bowen, 2015 (Bowen et al., 2015) | 1) Bowen et al. 2012# (Bowen et al., 2012)  
2) Weinstock et al. 2007+ (Weinstock et al., 2007)  
3) Weinstock et al. 2004** (Weinstock et al., 2004) |
| Geller, 2006, (Geller et al., 2006) | None provided |
| Glanz, 2010, (Glanz et al., 2010) | None provided |
| Glanz, 2013 (Glanz et al., 2013) | 4) Glanz et al. 2003# (Glanz et al., 2003)  
5) Glanz et al. 2010** (Glanz et al., 2010)  
6) Nelson et al. 2004# (Nelson et al., 2004) |
| Glanz, 2015 (Glanz et al., 2015) | 7) Glanz et al. 2010** (Glanz et al., 2010) |
| Glazebrook, 2006 (Glazebrook et al., 2006) | None provided |
| Heckman, 2016 (Heckman et al., 2016) | 8) Glanz et al. 2008** (Glanz et al., 2008) |
| Manne, 2010 (Manne et al., 2010) | 9) Manne et al., 2004# (Manne et al., 2004)  
11) Oliveria et al., 1999+ (Oliveria et al., 1999)  
12) Manne 2009# (Manne et al., 2009)  
13) American Academy of Dermatology (AAD), 2006 |
| Oliveria, 2004 (Oliveria et al., 2004) | 13) Weinstock et al. 1999+ (Weinstock et al., 1999)  
14) Weinstock et al. 2004** (Weinstock et al., 2004) |
| Rat, 2014 (Rat et al., 2014) | 15) Glanz et al. 2008** (Glanz et al., 2008) |
(www.who.int/uv/publications/en/UVIGuide.pdf) |
| Robinson, 2010 (Robinson et al., 2010) | 17) Robinson et al. 1998** (Robinson et al., 1998)  
18) Robinson et al. 2002** (Robinson et al., 2002) |
| Robinson, 2007 (Robinson et al., 2007a) | 19) Robinson et al. 2007** (Robinson et al., 2007a)  
20) Robinson et al. 2008+ (Robinson, Stapleton, & Turrisi, 2008)  
21) Robinson et al. 1998** (Robinson et al., 1998)  
22) Robinson et al. 2002** (Robinson et al., 2002) |
<table>
<thead>
<tr>
<th>Source</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turrisi, 2015 (Turrisi et al., 2015)</td>
<td>23) Boone et al. 2009* (Boone et al., 2009)</td>
</tr>
<tr>
<td></td>
<td>24) Robinson, Turrisi, Stapleton, et al., 2007** (Robinson, Turrisi, &amp; Stapleton, 2007b)</td>
</tr>
<tr>
<td></td>
<td>25) Robinson, Turrisi, and Stapleton, 2007** (Robinson et al., 2007a)</td>
</tr>
</tbody>
</table>

Note.
*Duplicate source cited in more than one trial report.
** Source coincides with one of the 13 trial reports included in this review.
# Not about SSE
+ Study reports on predictors or correlates of SSE and/or study aim was not to validate SSE measure / scale