Supplementary material

Design of knowledge assessment and assessment of content validity: The initial version of the knowledge assessment instrument contained 40 items, and expert validation was initially conducted by a cross-sectional representation of staff at both study sites based on job category and gender (Supplementary Table 1). Staff were requested to assess the relevance of each assessment based upon the following: 1) content relevance; 2) accuracy; 3) clarity; 4) breadth; and 5) appropriateness of format for knowledge transmission. Each item was assessed utilizing a 5-point score corresponding to: 1-High, 2-Good, 3-Medium, 4-Low, and 5-Very Low. To assess content validity, we utilized the content validation ratio (CVR), which deems each item as "essential," "useful", or "not necessary" (Supplementary Table 2)[38]. Of the five content relevance assessment items, the CVR identified content relevance and accuracy as the most important items.

A total of 12 staff, equally divided between the two sites and genders, evaluated the content validity of the assessments. The final assessment included 25 items after consideration of elimination, based upon their medical appropriateness, of those items with lower CVRs. In addition, the video script included 29 educational statements and used the same content evaluation scales. After development, the assessment was pilot tested (see appendix).

Pilot-testing: We pilot tested the knowledge assessment's 25 items on 20 individuals who were representative of each sites' demographics including four males (three Caucasians and one African-American or other race) and six females (five Caucasians

and one African-American or other race). Clarity was deemed the most important item attribute. For the final questionnaire, we retained all 25 items as more than 50% of respondents rated each item as essential.

Sample size calculation: We next performed a power analysis in order to determine the appropriate sample size pair for the matched case-control study. We utilized McNemar's test with the Type I error rate set at 5%. To achieve 80% power, the minimum sample size necessary to detect a statistical significance of 60% discordant pairs among all pairs is 89, when 5% of the discordant pairs are the situations where cases are in the intervention group but not control group[57].

Model description: Let Y_{it} indicate the test score of the ith subject at visit t. Here t_1,t_2 indicates the post-education and one-month follow-up visit respectively. The scores are assumed to follow a Poisson distribution, and they are modelled using a generalized linear mixed-effects model using the logarithmic (with natural basis) link function. The fix effects include gender, race, educational intervention, time point, and intervention by time point interaction, the pre-education test score as the baseline value and age. The inclusion of the pre-education test score controls for differences, for example, in education levels between the two sites. The random effect includes a random intercept γ_{0i}, which indicates the inter-individual variability among participants.

The model can be written as:

$$\ln(E[Y_{it}]) = \beta_0 + \beta_1 X_{age,i} + \beta_2 X_{gender,i} + \beta_3 X_{race,i} + \beta_4 X_{BL,i} + \beta_5 X_{trt,i} + \beta_6 X_{visit,it} + \beta_7 \left(X_{trt,i} \times X_{visit,it}\right) + \gamma_{0i} + \varepsilon_{it},$$

where $\gamma_{0i} \sim N(0, \sigma_0^2)$, $\varepsilon_{it} \sim N(0, \sigma^2)$, $i = 1,2,\ldots,176$, $t = 1,2$.

In the model, $X_{age,i}$, $X_{gender,i}$, $X_{race,i}$ represents the participant's age, gender and race; $X_{BL,i}$ is the baseline value of i^{th} participant, which uses the pre-education score; $X_{trt,i}$ indicates if the participant is in the control group (i.e. receives the brochure) or the intervention group (i.e. receives the video); $X_{visit,it}$ shows that if the i^{th} participant is in the post-education visit or the one-month follow-up visit and $X_{trt,i} \times X_{visit,it}$ is a interaction term of the intervention group and time point.

Weblink to video is available at:

https://drive.google.com/file/d/1mrJbXpRl7YrqbhoPH3sUphOTz2TNgc0e/view?usp=sharing

Supplementary Table 1: Position title and gender distribution of study site staff

Site	A			B		
	Position	Gender	#	Position	Gender	#
	Physician	Female	1	Physician	Female	1
		Male	2		Male	1
	RN	Female	1	RN	Female	1
	LPN	Male	1	LPN	Female	1
	LCSW	Female	1	LCSW	Male	2

assessing content validity.

Abbreviations: RN, registered nurse; LPN, licensed practical nurse; LCSW, licensed clinical social worker

Supplementary Table 2: Criteria used to assess the importance of each questionnaire item during pilot testing.

	Dimensions of Pilot-Testing		
Scale	Clarity	Personal relevance	Willingness to seek treatment
1	High	High	High
2	Good	Good	Good
3	Medium	Medium	Medium
4	Low	Low	Low
5	Very Low	Very Low	Very Low

Supplementary Table 3: Proposed and actual participant recruitment stratified by demographics and intervention time point.

Demographics			Proposed	Brochure		Video	
Race	Gender	Age	n	n (pre- and post-educational)	n (one-month follow-up)	n (pre- and post-educational)	n (one-month follow-up)
White	Female	≤ 39	21	21	19	19	17
White	Female	> 39	23	23	21	23	23
Other race	Female	≤ 39	4	4	4	4	3
Other race	Female	> 39	7	7	7	5	4

White	Male	≤ 39	17	17	14	17	16
White	Male	> 39	15	15	15	15	15
Other races	Male	≤ 39	2	2	2	2	2
Other race	Male	> 39	1	1	1	1	1
	Total		90	90	83	86	81

Abbreviations: IQR, interquartile ratio; SD, standard deviation