Online social integration is associated with reduced mortality risk

William R. Hobbs[a,b,c,1], Moira Burke[d], Nicholas A. Christakis[e,f,g,h], and James H. Fowler[a,i,1]

[a]Division of Social Sciences, University of California, San Diego, CA 92093; [b]Network Science Institute, Northeastern University, Boston, MA 02115; [c]Institute for Quantitative Social Science, Harvard University, Cambridge, MA 02138; [d]Facebook, Menlo Park, CA 94025; [e]Department of Sociology, Yale University, New Haven, CT 06520; [f]Department of Medicine, Yale University, New Haven, CT 06520; [g]Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520; [h]Department of Biomedical Engineering, Yale University, New Haven, CT 06520; and [i]School of Medicine, University of California, San Diego, CA 92093

Edited by Kenneth W. Wachter, University of California, Berkeley, CA, and approved September 15, 2016 (received for review April 6, 2016)

Social interactions increasingly take place online. Friendships and other offline social ties have been repeatedly associated with human longevity, but online interactions might have different properties. Here, we reference 12 million social media profiles against California Department of Public Health vital records and use longitudinal statistical models to assess whether social media use is associated with longer life. The results show that receiving requests to connect as friends online is associated with reduced mortality but initiating friendships is not. Additionally, online behaviors that indicate face-to-face social activity (like posting photos) are associated with reduced mortality, but online-only behaviors (like sending messages) have a nonlinear relationship, where moderate use is associated with the lowest mortality. These results suggest that online social integration is linked to lower risk for a wide variety of critical health problems. Although this is an associational study, it may be an important step in understanding how, on a global scale, online social networks might be adapted to improve modern populations' social and physical health.

social networks | social media | health | longevity | social support

People with more friends and more social ties in their community tend to live longer (1–4). Many researchers interpret this association as evidence that greater social support and social network integration lead to better health outcomes (4). For example, social integration is thought to improve health by motivating engagement in healthy behaviors (5, 6), improving immunity (7), and reducing inflammation (8). However, nearly all of this work has been conducted in the context of real-world, face-to-face social interactions. As more and more people use online social media to maintain friendships (as of June 2016, about 1.1 billion people use Facebook daily), an open question is whether or not this new context can be used to measure real world social activity and, distinctly, whether online social interactions are similarly associated with better health and increased human longevity.

Many researchers have shown that online access to friends promotes real world social activities (9). This finding suggests that online social media may increase the amount of overall social integration. To the extent that people use social media to coordinate and engage in healthy face-to-face social behavior, we might therefore expect a positive relationship between online use and health. However, it may also be the case that spending time on social media reduces the amount of time available for offline socializing. If so, then social media use might be more like watching television, which tends to crowd out social activity and, along with other forms of sedentary behavior, has been associated with worse health outcomes (10). Of course, because social media use is not randomly assigned, use might also be a proxy for other unmeasured traits.

To help adjudicate whether online social media use has a positive or negative relationship with health, we test here the associations between online social connection and human mortality. Using several different indicators, we measured deidentified counts of six-months' online social media activity of 12 million people on the website Facebook and evaluated whether these measures were associated with decreased mortality risk in a two year follow-up.

Results

Before analyzing online social connection and social media behavior, we compare mortality rates for the Facebook population vs. the population-at-large. In these analyses, we control for age and gender differences between the two groups, as well as a coarse proxy for race/ethnicity [based on data from the US Census Bureau (11); results are shown in *SI Appendix*] to account for known health disparities and slightly different levels of Facebook use by race/ethnicity. The age- and gender-matched mortality rate for the "full" population of Facebook users (*Materials and Methods*) was 63% of the rate in the California voter record (our data-matching benchmark). This association may result from difficulties in matching Facebook users to vital records. To more confidently evaluate the relative health of Facebook users compared with the general population, we focus our analysis on the "voter" subpopulation, which includes only those Facebook users also present in the California voter record. That is, we compare voters who are on Facebook to those who are not. The age- and gender-matched mortality rate for Facebook users within the voter record population was ∼88% of Facebook nonusers within the voter record population. In other words, the risk of dying in a given year is about 12% less for Facebook users than non-Facebook users.

We disaggregate this comparison by cause of mortality. Mortality due to sexually transmitted diseases, several types of cancer, unintentional injuries, drug overdoses, and suicides did not

Significance

People who have stronger social networks live longer. However, can we say the same about online social networks? Here, we conduct such a study. Using public California vital records, we compare 12 million Facebook users to nonusers. More importantly, we also look within Facebook users to explore how online social interactions—reflecting both online and offline social activity—are associated with longevity. We find that Facebook users who accept more friendships have a lower risk of mortality, but there is no relationship for those who initiate more friendships. Mortality risk is lowest for those with high levels of offline social interaction and moderate levels of online social interaction.

Author contributions: W.R.H., M.B., N.A.C., and J.H.F. designed research; W.R.H. analyzed data; and W.R.H., M.B., and J.H.F. wrote the paper.

Conflict of interest statement: M.B. is a Facebook employee. W.R.H. was a Facebook research intern in 2013.

This article is a PNAS Direct Submission.

Data deposition: We have created a dataset that conforms to the agency's guidelines for working with small cell sizes (20), and we will make this available to researchers who request it from the corresponding author.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. Email: whobbs@ucsd.edu or fowler@ucsd.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1605554113/-/DCSupplemental.

significantly differ between Facebook users and nonusers in the voter record (these results are shown in *SI Appendix*, Fig. 4). However, mortality due to infections [relative risk: 0.72; 95% confidence interval (CI): 0.63–0.82], diabetes (relative risk: 0.62; 95% CI: 0.56–0.70), mental illness or dementia (relative risk: 0.75; 95% CI: 0.67–0.83), ischemic heart disease (relative risk: 0.81; 95% CI: 0.76–0.86), stroke (relative risk: 0.71; 95% CI: 0.63–0.80), other cardiovascular diseases (relative risk: 0.88; 95% CI: 0.82–0.94), liver disease (relative risk: 0.65; 95% CI: 0.59–0.72), and homicide (relative risk: 0.55; 95% CI: 0.46–0.67) were all significantly lower for Facebook users than nonusers. Each of these associations remain significant with a Bonferroni correction for 17 comparisons. A coarse proxy for race/ethnicity (*SI Appendix*) strengthened the association between Facebook use and decreased mortality risk (relative risk: 0.86; 95% CI: 0.85–0.88).

It is important not to read too much into the comparison between Facebook users and nonusers because many factors may confound the apparent association between being a Facebook user and experiencing lower mortality. This is an observational result, and we have few socioeconomic controls because we do not have much information about nonusers. We cannot rule out the possibility that some seriously ill individuals signed up for Facebook to update friends on their condition or that Facebook might attract healthier individuals for reasons unrelated to their social connectedness. However, in the analyses that follow, we can make better inferences because they will be based on comparisons within the population of Facebook users where we can control for age, gender, marital status, device used to access Facebook, and sign-up date on Facebook, as well as friends' highest education and a coarse proxy for race/ethnicity (shown in *SI Appendix*).

Within the verified social media user population (*Materials and Methods*), we first analyze variation in use among Facebook users to explore what relationship different kinds of activities have with mortality. Quantity of social contacts has been associated with reduced mortality in a wide number of studies (4), but it is unclear what drives this association, and it has never been tested using online social connections. We separately analyzed the association between mortality and (*i*) initiated friendships (the user asks another user to be friends and that friend accepts) and (*ii*) accepted friendships (the user agrees to be friends with another user who asked). If online relationships are beneficial or otherwise predictive of good health, we would expect the number of accepted friendships to be associated with lower mortality risk. Additionally, if seeking social support is beneficial for one's health (or if it is associated with other beneficial personal attributes like capacity for self-care), we would expect initiated friendships to be associated with reduced mortality.

Fig. 1 plots all-cause mortality hazard ratio estimates and CIs by decile for the Facebook friending behaviors. Notice that accepted friendship requests were associated with lower mortality. In fact, the mortality rate for users with the most accepted friendships (highest decile) was about 66% (95% CI: 60–73%) of the rate for those with the least accepted friendships (lowest decile). However, there was no such linear association between mortality and sent friendship requests. These results replicate the classic relationship between reduced mortality and number of social contacts in a large scale online setting, but they suggest that what matters is not the tendency to seek out friends—it is the willingness of others to seek out and establish these friendships. To the extent that these results might be explained by some causal relationship between social support and health, the results suggest that merely seeking additional support may be ineffective.

We next consider whether Facebook interactions that are plausibly related to offline social activity are driving the relationship between increased social activity on Facebook and decreased mortality. Users who post status updates may also post photos in these updates, and many photos show face-to-face social interactions. In fact, past research suggests that tags in photos are a strong predictor that two people have a face-to-face relationship (12). In contrast, text-based interactions are less predictive (in fact, in a model including several types of interactions, "likes" were actually negatively associated with likelihood of an offline relationship) (12). In Fig. 2, we show all-cause mortality hazard ratio estimates for various combinations of text-based activity (as measured by posting statuses) and photo-based activity (as measured by posting photos). Notice that mortality risk declines with increased photos, whereas it actually increases with increased statuses. We replicate these results in a proportional hazard model that includes both of these measures (*SI Appendix*). Mortality risk was about 70% of average (95% CI: 54–91%) for those who post many photos (highest decile) but few statuses (lowest decile). These results are suggestive that offline social activities—and not online activities—are driving the relationship between overall Facebook activity and decreased mortality risk.

Posting statuses and photos are both undirected activities because they do not have a specifically intended recipient, so we wanted to explore the extent to which we would find different

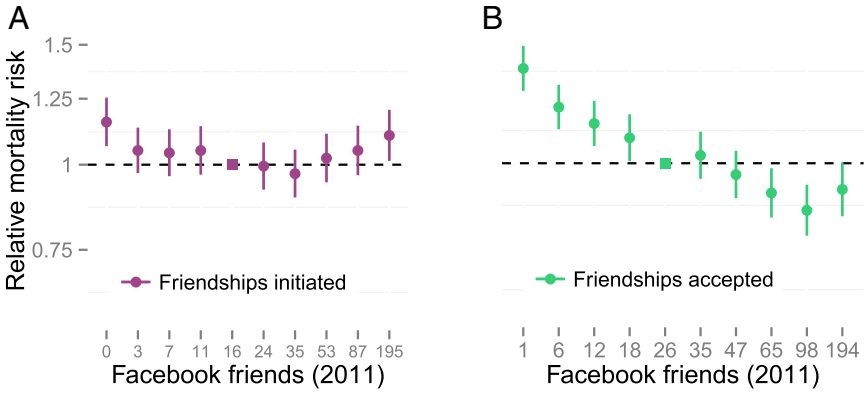

Fig. 1. Facebook friends and relative mortality risk (all-cause mortality). This figure shows all-cause mortality estimates (points) by deciles of Facebook friend counts, by initiated (*A*) and accepted (*B*) Facebook friendships and adjusted for age, gender, device use, and length of time on Facebook. The vertical bars are 95% CIs, and the square at 1.0 is the reference category. (*A*) Initiated friendship: the subject sent a Facebook friendship request that was then accepted. (*B*) Accepted friendship: the subject received and accepted a friendship request. The *x* axis is the median number of Facebook friends in the decile, and the *y* axis is the relative mortality risk estimated in a Cox proportional hazard model. The two friendship categories were estimated separately due to high collinearity between them.

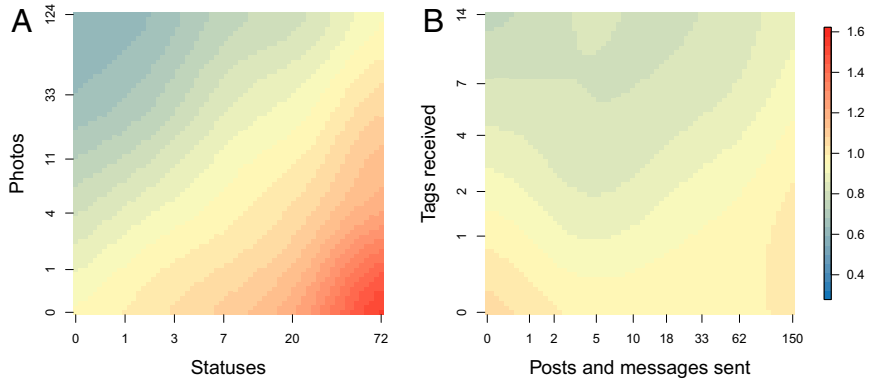

Fig. 2. Mortality risk as a function of two social media activities. (*A*) All-cause mortality hazard ratio estimates for combinations of text-based activity (as measured by posting statuses) and photo-based activity (as measured by posting photos). (*B*) All-cause mortality hazard ratio estimates for combinations of text-based directed communications (as measured by posts with tags and messages sent) and photo-based directed communications (as measured by photo tags received). The *x* and *y* axes are the quantiles of activities shown in the labels (each tick shows the respective quantile median). Colors show risk of mortality at each combination (red is higher; blue is lower). Results were estimated in a Cox proportional hazard model with interacted indicators for each decile of activity (fewer than 10 categories where the number of users with 0 activity spanned more than one decile). To present similar relative hazard scales (where a hazard of 1.0 corresponds to an average risk), we used the [0,0] interaction quantile as the reference category in the undirected analysis (no photos, no statuses) and the [1, 1] interaction quantile as the reference category in the directed analysis (one sent post or message, one received photo tag). We smoothed the quantile estimates with an approximate Nadaraya–Watson kernel smoother set to a bandwidth of 0.9.

relationships in directed communications on Facebook. Previous research suggests directed communication online is linked to increases in perceived social support, while undirected communication is not (13). Fig. 2 shows all-cause mortality hazard ratio estimates for text-based directed communications (as measured by posts with tags and messages sent) versus photo-based directed communications (as measured by photo tags received). We focus on received tags rather than sent tags for photo-based interactions because they usually indicate that the recipient is in the photo and therefore engaging in a real-world social activity (for comparisons of sent and received tags, see *SI Appendix*). Here, we find, once again, an inverse relationship between photo-based activity and mortality risk. However, there is now a nonlinear relationship with online-directed communications: extremely high and extremely low levels of online-directed communication are both associated with higher mortality. We replicate these results in a proportional hazard model that includes both of these measures and an additional nonlinear term for the online communication measure (*SI Appendix*). This model suggests that mortality risk is lowest for those who are tagged in many photos (highest decile) and who engage in a moderate level of online-directed communication (fourth to sixth deciles).

Last, we examine whether online social activities more strongly predict mortality due to causes that are more likely to be related to social factors. We present cause-specific estimates in order, from least expected to be predicted by social support to most expected. Social support and integration studies tend to find stronger relationships between social support and cardiovascular-caused mortality than cancer-caused mortality (4), and past work on substance abuse and suicide leads us to expect the largest effects for these causes (14–16).

Fig. 3*A* shows that the number of online friendships is not significantly related to decreased mortality due to cancer but is for cardiovascular disease (91%; 95% CI: 87–96%) and even more so for drug overdose (78%; 95% CI: 70–87%) and suicide (73%; 95% CI: 66–80%). Moreover, when we separately analyze initiated and accepted friendships, the results suggest that accepted friendships are driving the overall relationship, as we previously showed in Fig. 1.

Fig. 3*B* shows that increased photo uploads is associated with reduced mortality for all major causes except suicide. In contrast, there is a small positive relationship between status updates and cancer, perhaps reflecting an effort by Facebook users to seek social support or broadcast updates online once they are diagnosed with a chronic illness. Status updates appear to be uncorrelated with mortality for drug overdoses and suicides.

Finally, Fig. 3*C* shows that sent text-based communications are generally unrelated to mortality risk for these causes, but received communications tend to predict higher risk of mortality due to cancer (108%; 95% CI: 104–112%) and lower risk due to drug overdose (88%; 95% CI: 80–96%) and suicide (82%; 95% CI: 74–90%). Once again, this association suggests that social media is being used by cancer victims to broadcast updates, which elicit received messages, and the contrast between cancer (a positive relationship) and other causes (a negative relationship) may help to explain the nonlinear relationship observed with all-cause mortality in Fig. 2. Meanwhile, received photo tags, our strongest indicator of real-world social activity, are strongly inversely associated with all types of mortality except those due to cancer, and the inverse relationship is strongest with drug overdose (70%; 95% CI: 64–77%) and suicide (69%; 95% CI: 63–76%).

All of the relationships in Fig. 3 remain significant with a Bonferroni correction for 11 comparisons, with the exception of (*i*) drug overdose and friendships initiated, (*ii*) drug overdose and sent communications, and (*iii*) suicide and sent communications.

Discussion

This large-scale study measured the extent to which online social life predicts differences in mortality rates, confirming that it does so in the same way that offline social life does. Our results suggest that people who use online social media experience lower mortality rates than those who do not. Although this finding is very likely to be explained at least in part by differences in socioeconomic status between those who have access to social media and those who do not, we show that among a higher status group (registered voters), this difference in mortality persists. To the extent that online social media platforms like Facebook provide an opportunity to maintain social relationships, they may also indirectly provide people with greater capacity to receive social support and encourage socially-motivated behaviors that may prevent illness.

Among those who do use social media, overall network size is associated with better health. Just like numerous past studies of real world social networks, we find that people with more friends online are less likely to die than their disconnected counterparts. This evidence contradicts assertions that social media have had a

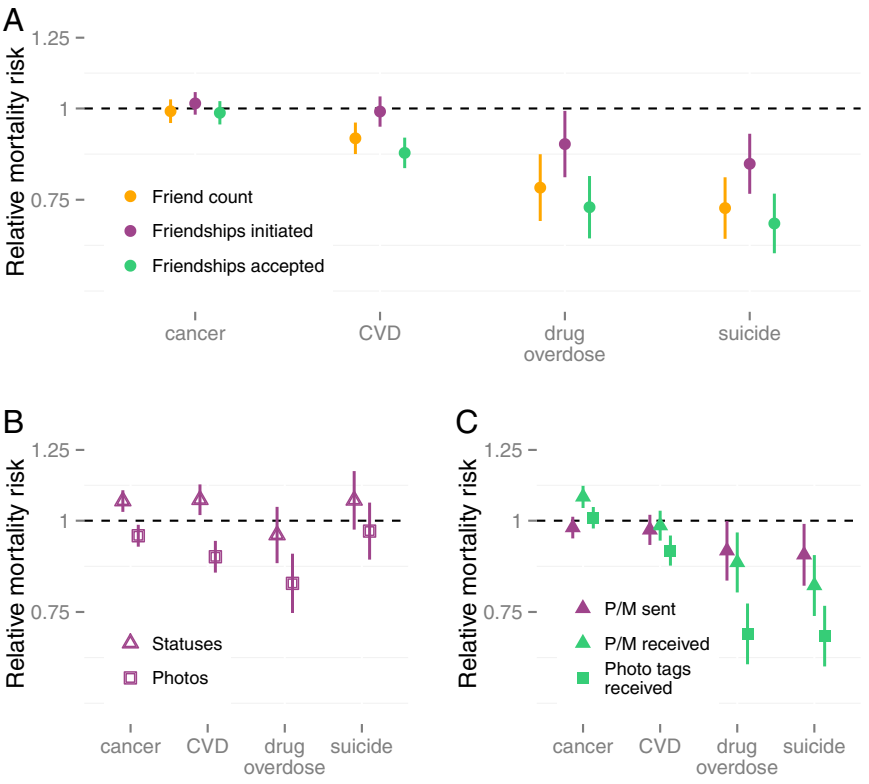

Fig. 3. Cause-specific mortality risk as a function of various Facebook activities. (*A*) Cause-specific mortality hazard ratio estimates for Facebook friendships (total, initiated, and accepted counts). (*B*) Cause-specific mortality hazard ratio estimates for undirected activities (social media communications without a specific recipient). (*C*) Cause-specific mortality hazard ratio estimates for directed activities (social media communications with a specific recipient). Points indicate cause-specific morality risk estimates and vertical bars show 95% CIs. Actions by the subject are shown in purple, actions by the subject's friend are shown in green, combined actions are shown in orange, circles denote friending activities, triangles denote text-based actions, and squares denote photo-based actions. All variables are logged (after adding one to the activity count), scaled by their SD (so that a unit change is an SD and comparable across activities), and centered at their means. The *x* axis is the cause of death (an extended table for more and more finely discriminated causes of death is included in *SI Appendix*), and the *y* axis is the relative mortality risk associated with a SD change in the relevant activity, estimated in a Cox proportional hazard model. All activity categories were estimated separately due to high collinearity among them. "P/M" here is an abbreviation for "posts and messages."

net-negative impact on health. Both comparisons between users and nonusers and between low users and high users suggest that social media use is predictive of lower mortality. If social media use were extremely unhealthy, we would expect to find an overall positive relationship between use and mortality, but we do not.

Moreover, those measures of online social activity that are most predictive of reduced mortality are precisely those that are most likely to promote or otherwise indicate offline social interactions. In fact, when we separately analyze activities like posting statuses and posting photos, we find that, underneath the overall relationship between use and reduced mortality, there is an opposite association with behaviors that are likely to be disconnected from the offline world.

Notably, the relationships between offline and online social activity vary by underlying cause of mortality in informative and theoretically consistent manners. Cardiovascular disease is more strongly related to social factors than cancer, and associations between cardiovascular disease and social isolation are stronger in offline interactions than in online ones. Online social interactions and social support most strongly predict better health for underlying causes of death related to mental illness and substance abuse (i.e., drug overdose and suicide), where we expect the largest social support effects.

A major advantage of using online data are that we can distinguish between friendships sought and friendships accepted, and our results show that the relationship between friendship and reduced mortality is driven by others' perception of closeness and desire to connect online. These results suggest that better health is unlikely to be determined solely by an individual's ability to meet more people or to seek out connection. Instead, it depends on the likelihood that, once having met, social interactions will continue, and/or others will perceive and maintain the friendship. Unfortunately, the finding suggests that interventions that try to increase capacity to seek support may not have the intended effect of improving health.

This study has many limitations. For one, our study may have limited external validity because Facebook is unique among social media sites, and online platforms are constantly changing; by analogy, efforts to predict real world flu using online searches failed when the algorithms were not updated to reflect changing use patterns (17). Another concern is that our measures of mortality risk cover just 2 y for a single state in the United States (California). It is possible that we might find different relationships in studies with longer follow-up or in different states or countries. However, the most notable limitation involves the classic difficulty in distinguishing association from causation that limits all observational studies. Although we show many relationships between social media use and reduced mortality risk, we have not provided evidence of a causal relationship here.

We cannot say that spurring users to post more photos on Facebook will increase user longevity. On the other hand, observational studies are often an important first step for better understanding new phenomena. We hope this study plays a role in spurring interest in online social effects on health, just as Lisa Berkman and Leonard Syme's classic paper in 1979 spurred interest in the relationship between social support and longevity (1).

In particular, the observational relationships reported here suggest that online applications that explicitly promote offline social life may generate positive health outcomes. Indeed, some scholars are currently designing, implementing, and testing such applications (18), and we encourage others to do so as well.

Materials and Methods

To conduct this study, we needed access to Facebook data and also to publicly available vital records. Our study protocol was approved by three different bodies: the institutional review board at the University of California, San Diego (UCSD); the State of California Committee for the Protection of Human Subjects; and the Vital Statistics Advisory Committee at the California Department of Public Health. In addition, the research was reviewed by Facebook's internal review group. The UCSD institutional review board approved a waiver of informed consent for analysis of existing data.

We worked directly with Facebook to analyze aggregate counts of use data. We restricted our analysis to California where we could also access public vital records containing cause of mortality. We identified all California-based Facebook users who (*i*) signed up before October 2010, so they would have at least 90 d of experience with the site before January 2011, when we started measuring social media use; (*ii*) listed a first name (or nickname), last name, and date of birth not shared by any other person in California, so that these individuals could be linked to other records uniquely; and (*iii*) were born between 1945 and 1989, because use at the time was rare among people with earlier birth years, and some people mistakenly reported their birth year as 1990 because it was the default when joining Facebook. We also required first names and last names to appear at least once (independently) in the California voter record, and we omitted users with a birthday of January 1 because it was the default; 12,689,047 profiles fit these criteria (the full population), of whom 4,011,852 were also present in the California voter record (by complete first name, last name, and date of birth: the voter subpopulation). To preserve privacy, after automatically matching to public records, all analyses were performed on deidentified, aggregate data. All data were observational; no one's experience on the site was different from usual.

Once we identified the eligible population, we compared user information (first name or nickname, last name, and date of birth) to California Department of Public Health vital records for 2012 and 2013 to ascertain mortality status and cause of mortality. For the period of study, the Facebook and California voter record populations differed in their age and gender distributions because many people on Facebook tend to be younger than the population of California voters. To ensure age and gender covariate balance in our analyses, we compared all deceased individuals on Facebook to a stratified random sample of nondeceased individuals (*SI Appendix*, Fig. 5) from the full and voter populations described above. There were 179,345 people in our age- and gender-based probability sample of Facebook users born between 1945 and 1989, of whom 17,990 died between January 2012 and December 2013; 89,597 were also present in the California voter record, of whom 11,995 had died between January 2012 and December 2013.

We categorized underlying causes of mortality in 17 specific categories, as well as 4 broader categories (cancer, cardiovascular disease, drug overdose, suicide), based on codes of the *International Classification of Diseases, Tenth Revision* (19). We categorized methods of interaction on Facebook using two basic dichotomies: directed (e.g., messages) versus undirected (e.g., status updates) and text-based (e.g., wall posts) versus photo-based (e.g., photo tags). We further distinguished directed actions as outgoing (i.e., sent) or incoming (i.e., received). These activity categories corresponded to the major variance components of a principal component analysis of Facebook activities (shown in *SI Appendix*).

We used the Cox proportional hazard model to estimate relationships between Facebook activities and mortality. The start time was the user's age on January 1, 2012, and the end time was the user's age when deceased or as of December 31, 2013. We classified the failure event as 1 for deceased in "all-cause" models and 0 otherwise and 1 for mortality from specific causes (e.g., 1 for mortality caused by cancer) and 0 otherwise. All models are stratified by gender and, to account for possible socioeconomic status confounders, include controls for (*i*) Facebook sign-up date; (*ii*) access method most commonly used on the site (e.g., the www.facebook.com webpage, the mobile page m.facebook.com, or a native app on a smartphone); and (*iii*) whether the individual used a smartphone application to access Facebook social tools during the observation period. The additional socioeconomic status control "highest education levels listed by friends on Facebook" (shown in *SI Appendix*) did not substantively alter the results.

For the proportional hazard models, all Facebook use variables were logged [with $log(x+1)$ so that 0 values were well-defined] because online counts of behavior are skewed (most users have low activity and some users have very high activity). Intuitively, this transformation assumes that one additional photo tag (or any other action) is less important for good/bad health and social support once a user already has many of them. For other low, medium, and high use comparisons, we discretized variables by deciles, combining zero-activity deciles when they spanned multiple deciles. This discretization allows us to estimate nonlinear effects of Facebook use, as well as the effects of unusually high levels of online activity. We report details of all models in *SI Appendix*.

To compare the Facebook population to the California voter record population, we weighted the population in the California voter record to exactly resemble the distribution of age and gender on Facebook.

P values for all estimates are two-tailed. We computed CIs using robust SEs. Because we are more confident that the voter subsample represents real people, we show the results for this subsample in the main text, but we also show results for the full sample in *SI Appendix*.

The California Department of Public Health prohibits the release of individual-level data. However, the agency does allow the release of aggregated data that protects data confidentiality. We have created a dataset that conforms to the agency's guidelines for working with small cell sizes (20), and we will make this available to researchers who request it from the corresponding author.

ACKNOWLEDGMENTS. We thank Cameron Marlow, Lada Adamic, Danny Ferrante, Arturo Bejas, Pete Fleming, Will Nevius, and Molly Jackman for their support on this project.

1. Berkman LF, Syme SL (1979) Social networks, host resistance, and mortality: A nine-year follow-up study of Alameda County residents. *Am J Epidemiol* 109(2): 186–204.
2. House JS, Landis KR, Umberson D (1988) Social relationships and health. *Science* 241(4865):540–545.
3. Smith KP, Christakis NA (2008) Social networks and health. *Annu Rev Sociol* 34(1): 405–429.
4. Holt-Lunstad J, Smith TB, Layton JB (2010) Social relationships and mortality risk: A meta-analytic review. *PLoS Med* 7(7):e1000316.
5. Ford ES, Ahluwalia IB, Galuska DA (2000) Social relationships and cardiovascular disease risk factors: Findings from the third national health and nutrition examination survey. *Prev Med* 30(2):83–92.
6. Umberson D, Crosnoe R, Reczek C (2010) Social relationships and health behavior across life course. *Annu Rev Sociol* 36(1):139–157.
7. Cohen S, Doyle WJ, Skoner DP, Rabin BS, Gwaltney JM, Jr (1997) Social ties and susceptibility to the common cold. *JAMA* 277(24):1940–1944.
8. Loucks EB, Berkman LF, Gruenewald TL, Seeman TE (2006) Relation of social integration to inflammatory marker concentrations in men and women 70 to 79 years. *Am J Cardiol* 97(7):1010–1016.
9. Rainie H, Wellman B (2012) *Networked: The New Social Operating System* (MIT Press, Cambridge, MA).
10. Thorp AA, Owen N, Neuhaus M, Dunstan DW (2011) Sedentary behaviors and subsequent health outcomes in adults a systematic review of longitudinal studies, 1996-2011. *Am J Prev Med* 41(2):207–215.
11. US Census Bureau (2000) Genealogy data: Frequently occurring surnames from census 2000. Available at www.census.gov/topics/population/genealogy/data/2000_surnames.html. Accessed March 24, 2014.
12. Jones JJ, et al. (2013) Inferring tie strength from online directed behavior. *PLoS One* 8(1):e52168.
13. Burke M, Kraut R (2016) The relationship between Facebook use and well-being depends on communication type and tie strength. *J Comput-Mediat Comm* 21: 265–281.
14. Conwell Y, Duberstein PR, Caine ED (2002) Risk factors for suicide in later life. *Biol Psychiatry* 52(3):193–204.
15. Rossow I, Lauritzen G (1999) Balancing on the edge of death: Suicide attempts and life-threatening overdoses among drug addicts. *Addiction* 94(2):209–219.
16. Hawton K, Houston K, Haw C, Townsend E, Harriss L (2003) Comorbidity of axis I and axis II disorders in patients who attempted suicide. *Am J Psychiatry* 160(8):1494–1500.
17. Lazer D, Kennedy R, King G, Vespignani A (2014) Big data. The parable of Google Flu: Traps in big data analysis. *Science* 343(6176):1203–1205.
18. Patrick K, et al. (2014) Design and implementation of a randomized controlled social and mobile weight loss trial for young adults (project SMART). *Contemp Clin Trials* 37(1):10–18.
19. World Health Organization (1990) *International Statistical Classification of Diseases and Health Related Problems* (World Health Organization, Geneva), 10th Ed.
20. California Department of Public Health Office of AIDS (2014) *CDPH/OA Internal Guidelines for Working with Small Cell Sizes* (California Department of Public Health Office of AIDS, Sacramento, CA). Available at https://www.cdph.ca.gov/programs/aids/Documents/SmallCellGuidance.pdf. Accessed June 13, 2016.

Online Interaction, Social Support, and Health

William R. Hobbs, Moira K. Burke, Nicholas A. Christakis, James H. Fowler

A About Facebook 2

B Data linking and controls 2

C Sampling rates 5

D Activity categories 6

E Models included in Figure 4 9

F Models included in Figure 1 12

G Models included in Figure 2 14

H Continuous estimates for relationships in Figure 2 16

I Models included in Figure 3 18

J Extended cause-specific estimates for models included in Figure 3 26

K Results for full (non-voter included) population 29

L Results for 'voter' population with controls added for 'full' population analyses 34

M Results for 'voter' population with race/ethnicity controls 37

Glossary
Profile The user's personal page. Contains basic information and social activities. **Wall** A message board on the user's profile that is visible to friends. **Facebook friend** A connected person who can contact the user and see the user's profile and updates.
Friend request A request to become Facebook friends. **Wall post** A message written to a friend's wall that can be seen by other friends. **Message** A message sent to a friend or friends that can be seen only by the sender and recipient(s). **Photo** An uploaded photo, usually visible to all friends. **Photo tag** An added label associating a person with an uploaded photo. **Status** A broadcasted post on the user's own profile, usually visible to all friends. May contain a photo.
News feed A stream of updates of friends' activities.

Table 1

A. About Facebook

Facebook (*www.facebook.com*) is an online social networking website founded in 2004 and is used by about 71% of Americans [1]. The website offers a number of online tools to keep in touch with friends, including 'messages' (private directed communications seen only by the sender and recipient(s)), 'wall posts' (quasi-public directed communications posted to the recipient's Facebook 'wall' and visible to the recipient's friends), and 'status updates' (quasi-public undirected communications that are visible to the sender's friends who see them in a continuous stream of all their friends' updates called a 'news feed'). The website also offers tools to post photos in which one can 'tag' (label) friends (usually because they appear in the photos).

Each user maintains a 'profile' on the site. These profiles contain basic information such as name, education, friends, status updates, and location, as well as 'liked' 'pages' (usually interests, such as music and news sites). These profiles are further associated on the website with a 'timeline' history of all online activity and social interactions that have occurred since the user joined the site.

B. Data linking and controls

Once we identified the eligible population, we compared user information (first name or nickname, last name, and date of birth) to California Department of Public Health vital records for 2012 and 2013 to ascertain mortality status and cause of mortality. We then linked users who were living in January 2012 to their aggregated Facebook usage for the six-month period January 2011 through June 2011, as well as basic demographic information: year of birth, gender, date signed up on Facebook, highest education listed on profile, marital status listed on profile, and type of device used to access Facebook, along with the same information for all Facebook friends of the subjects. We excluded the six months prior to the impanel date (July to December 2011) so that the observation period was less likely to include acute periods of illness and disability. To be clear: we are testing associations between 1) social media usage over a six month period and 2) mortality over a subsequent 24 month period, with a 6 month gap between these two measurement periods.

Prior analyses of social media usage have typically restricted their focus to relatively active users (e.g. active on a specific day [2]) and/or self-reports of activity [3, 4]. The reason for this is that Facebook and other social media websites do not routinely verify that users have only one profile or are using their real names, except when necessary for operational reasons (e.g. spam or advertising audience estimates). However, low frequency use is an important variable of interest in the present analysis, so we do not select users based on minimal activity criteria.

Instead, we require only that individuals 1) had at least one Facebook friend as of January 2011 so that use of the website is at least minimally social; and 2) listed a first name (or, based on a publicly available database, associated nickname) and last name present in the California voter record (e.g. we included users named 'Jenny' if anyone in the California voter record was named 'Jenny' or 'Jennifer'). To further confirm that users listed real names, we segmented our analyses based on whether individuals listed a first name (or nickname), last name, and date of birth on Facebook that was also present in the California voter record, and we omitted users who listed a January 1st birthday because this is the default value when signing up for the site. 12,689,047 profiles fit these criteria (the 'full' population), of whom 4,011,852 were present in the California voter record (the 'voter' subpopulation). This match rate of 32% is similar to the match rate reported in previous analyses of California Facebook users [2] and is consistent with the observations that 1) younger people are less likely to be registered to vote, and 2) California has the 2nd lowest voter registration rate in the United States because of its large non-citizen population.

Cause of death categorizations differ from standard categorizations seen in prior works [5] only in that there are fewer old age categories (e.g. no dedicated prostate cancer category) and more young age categories (e.g. distinguishing between drug overdoses and unintentional injuries).

In Figure 4 we compare mortality rates on Facebook vs. in the general population by cause of mortality. Mortality due to sexually transmitted diseases, several types of cancer, unintentional injuries, drug overdoses, and suicides did not significantly differ between Facebook users and nonusers in the voter record. However, mortality due to infections (relative risk 0.72, 95% CI 0.63 to 0.82), diabetes (0.62, CI 0.56 to 0.70), mental illness or dementia (0.75, CI 0.67 to 0.83), ischemic heart disease (0.81, CI 0.76 to 0.86), stroke (0.71, CI 0.63 to 0.80), other cardiovascular diseases (0.88, CI 0.82 to 0.94), liver disease (0.65, CI 0.59 to 0.72), and homicide (0.55, CI 0.46 to 0.67) were all significantly lower for Facebook users than nonusers.

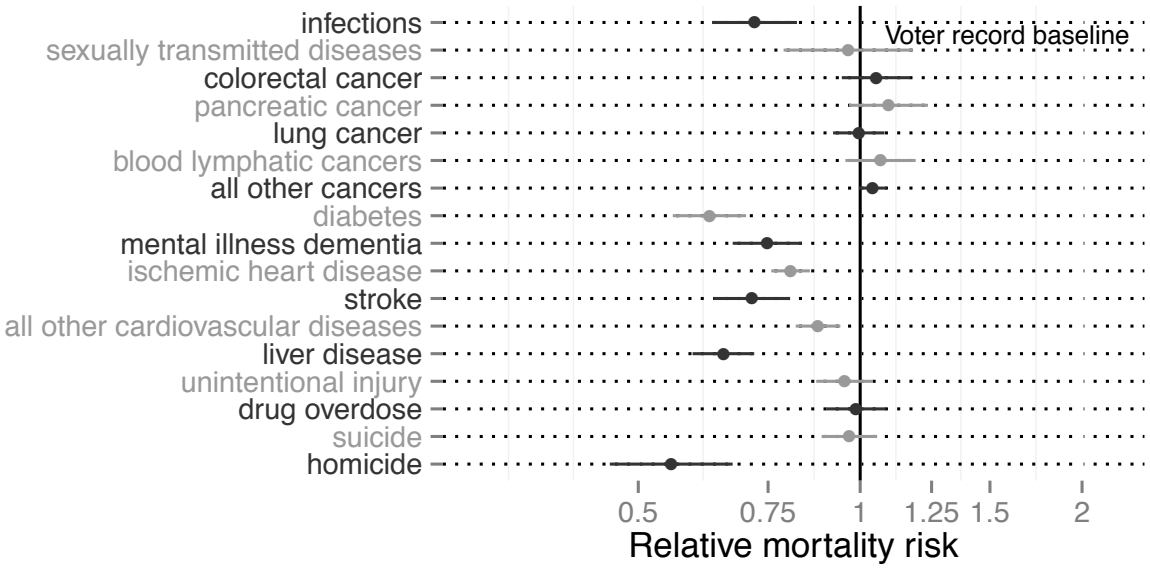

Fig. 4. *Facebook, California voter record mortality rate by cause comparison.* This figure shows relative cause-specific mortality risk among Facebook users whose first name, last name, and date of birth are listed in the California voter record compared to all California registered voters. The y-axis (on bottom) is the relative mortality risk estimated in a Cox proportional hazard model, and the x-axis is the cause of death. The all registered voter comparison group was exactly matched on age and gender so that the comparisons groups are, by-design, perfectly balanced on these covariates.

C. Sampling rates

For the period of study, the Facebook and California voter record populations differed substantially in their age, and to some extent gender, distributions, as many Facebook users were relatively young and many voters were relatively old. For Facebook activity models, we sampled California Facebook users at 10 times the age- and gender-specific California mortality rate (mortality rate interpolation shown in Figure 5).

The average age as of January 2011 was 49.6 (sd 12.4) and 42% listed female gender. Of those who made any social action on the website for January 2011 through June 2011, 22% used smartphone applications on iOS, Android, or Blackberry operating systems (for comparison, others have reported that 35% of Americans owned a smartphone in 2011 [6]). The median number of Facebook friends was 48 (25th percentile 18, 75th percentile 114; mean 108, sd 220) and the median number of individuals with whom a user interacted at least once during the study period was 26 (25th percentile 5, 75th percentile 79; mean 68, sd 137). These numbers are lower than those for all Facebook users, but note that social connections and social media activity are typically lower in older populations.

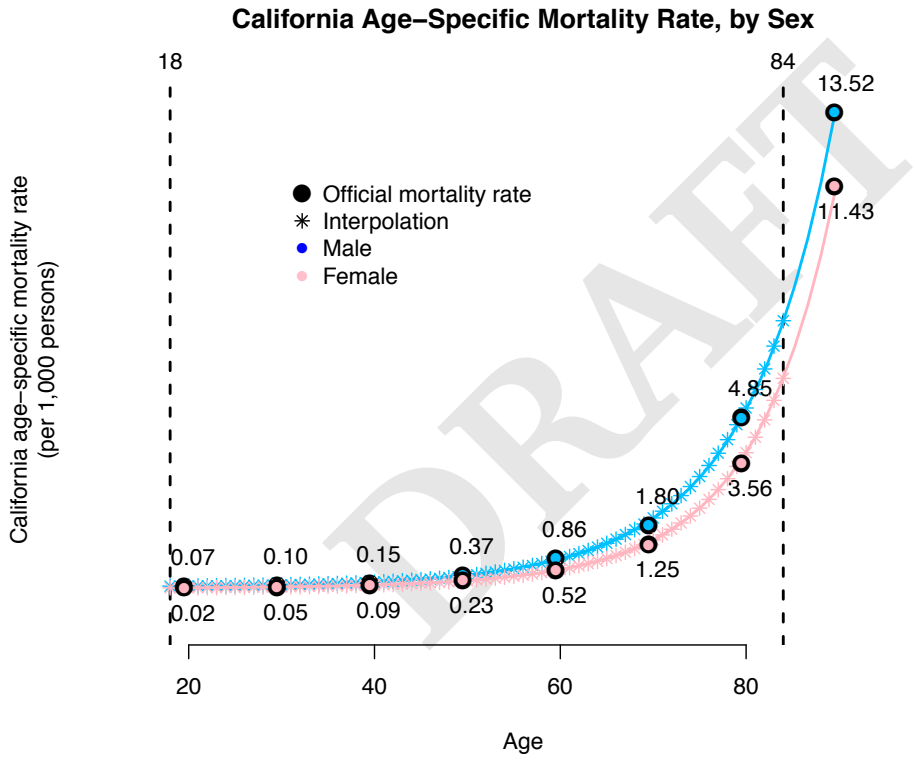

Fig. 5. *Sampling rates by age and gender.* This figure shows the sampling rates (interpolated mortality rates) used to construct an age and gender balanced sample for the Facebook-only analyses. California-based Facebook users were randomly sampled at a rate 10 times the California age and gender-specific mortality rate and included in a de-identified analysis file that included all decedents.

D. Activity categories

All Facebook activities were correlated because people who signed in more often were more likely to interact with each other using a variety of different methods. Correlations between activities ranged from 0.85 for sent versus received messages to 0.33 for posting statuses versus being tagged in photos. The full correlation matrix for major Facebook activities is in Supplementary Figure 6.

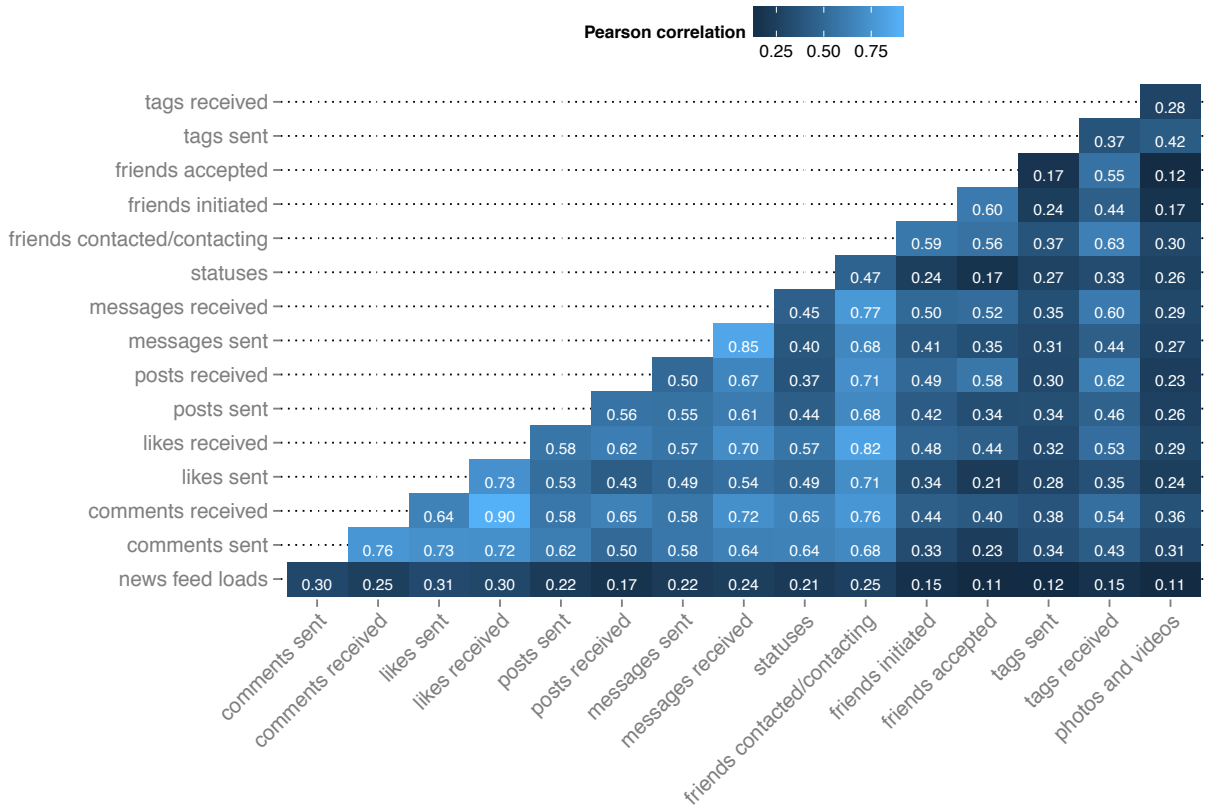

Fig. 6. *Correlations of common Facebook activities.* This figures shows the Pearson correlations of major Facebook activities. These variables are ordered by the 3rd component (online versus offline) in the principal component analysis shown in Figure 7.

In our longitudinal statistical analyses, we distinguished out-going (i.e. sent) from incoming (i.e. received) updates, communications, and requests for 'friendship' status on the website, as well as text-only (i.e. primarily online) and photo-based (i.e. more likely originally offline) social interactions. Because low levels of online social activity might reflect social isolation, while very high levels might replace physical, offline social activity or reflect a social network's provision of social support following news of terminal illness, we also compared low, moderate, and high usage associations.

These activity categories correspond to loadings on major variance dimensions in a principal component analysis [7], where overall activity is the first dimension, sent versus received messages is the second dimension, and text versus photo activity is the third dimension. Loadings (i.e. transformation coefficients) in Figure 7 are the eigenvectors of our sample's Facebook activity covariance matrix and they can be multiplied by the original activity counts to produce a transformation of the data which preserves correlated information in a smaller number of composite, orthogonal variables. The eigenvector corresponding to the leading eigenvalue contains each variable's contribution to the matrix's largest variance component (and the eigenvector correponding to the *n*th largest eigenvalue is each variable's contribution to the *n*th largest variance component). In the principal component analysis here, we $log(x+1)$ and scale each variable by its standard deviation and then center at zero.

The first component (i.e. the largest variance component, accounting for 70% of the proportion of variance explained) in this principal component analysis is overall activity (not shown because it is similar for all activity variables), the second is undirected/outgoing activity versus incoming activity (including network size – this component explains 7% of the variance), and the third is text-based versus photo-based activity (5% of the variance). In other words, it appears that users vary primarily in their level of activity, the extent to which the send or receive interactions, and the extent to which they use photos or text to interact with others on the site.

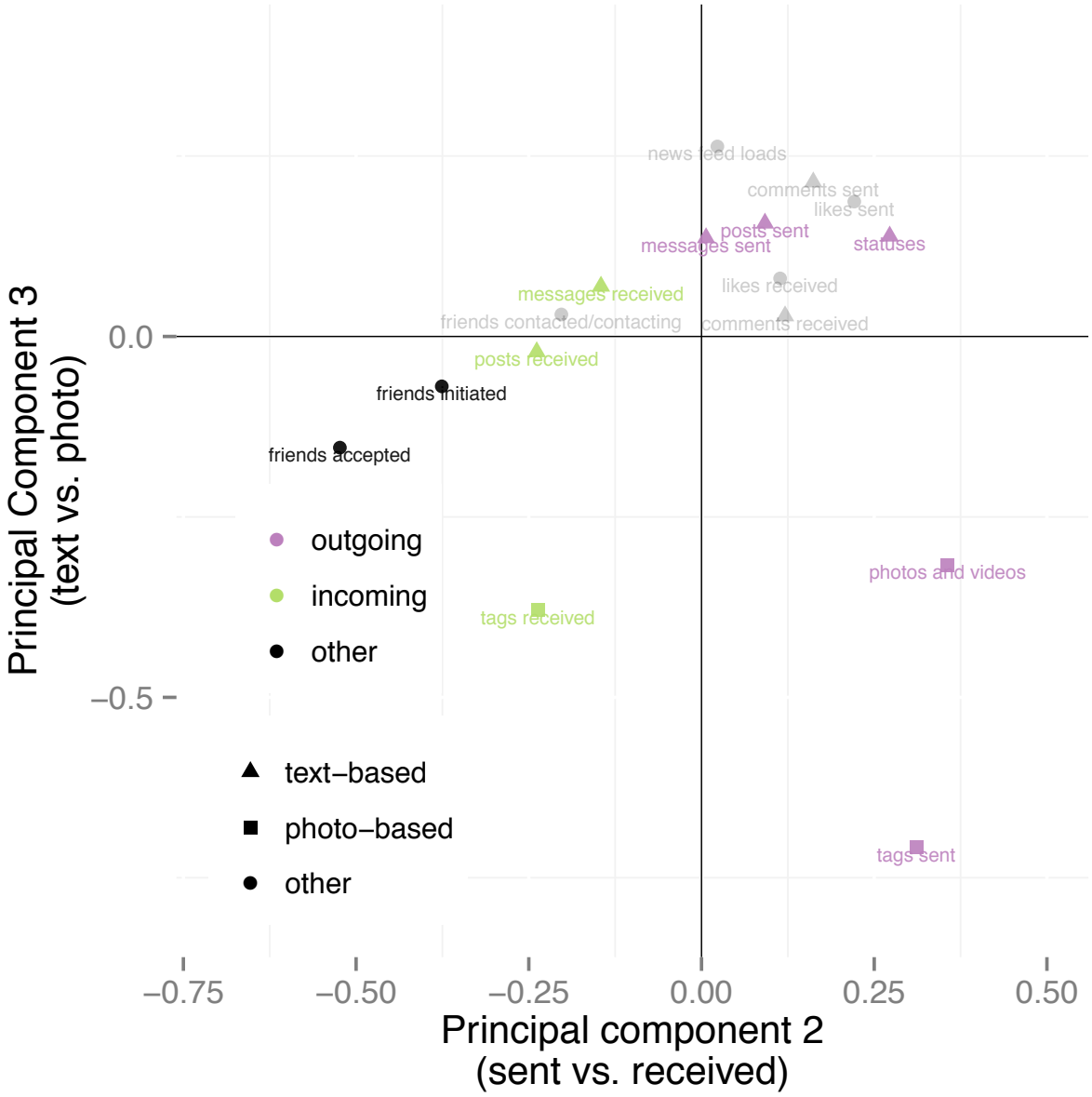

Fig. 7. *Principal component analysis of common Facebook activities.* This figure shows the second and third components of a principal component analysis on major Facebook activities. The first component (not shown) is the overall level of activity, and the activities of interest do not clearly differ on it. The second component corresponds to sent versus received messages, and the third component corresponds to text versus photo activities.

E. Models included in Figure 4

For Facebook population comparisons to the California voter record comparison populations, we matched Facebook and the voter record population by age and gender. We summed the number of individuals in each age-gender combination stratum on Facebook, and weighted the individuals in the California voter record to match the same age and gender distribution using the formula $m_t^s/m_c^s \times m_c/m_t$, where m_t^s is the number of observations in the Facebook strata and m_c^s is the number of observations in the voter record strata (m^s and m^c are the overall numbers of observations in each data set).

For the by-cause estimates in Figure 4, we use the Cox proportional hazard model, with added weights so that the datasets are matched exactly, and plot the estimate for an indicator variable signifying presence on Facebook. Because not all voters list their gender on voter registraton forms (and this is not an option on many forms), we relied on gendered first names in place of reported gender in this analysis (using the same last name gender proportion cutoffs for both Facebook and the voter record), and excluded ambiguous names (where less than 95% of reporting registrants listed a specific gender).

Table 2. Infections risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.719
	(0.630, 0.821)
Male	1.660
	(1.486, 1.855)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 3. Sexually transmitted diseases risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.963
	(0.787, 1.178)
Male	9.074
	(6.826, 12.062)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 4. Colorectal cancer risk, voter record.

	Dependent variable:
	Deceased
Has social media account	1.051
	(0.944, 1.170)
Male	1.274
	(1.156, 1.404)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 5. Pancreatic cancer risk, voter record.

	Dependent variable:
	Deceased
Has social media account	1.092
	(0.965, 1.236)
Male	1.475
	(1.317, 1.652)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 6. Lung cancer risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.996
	(0.919, 1.079)
Male	1.138
	(1.059, 1.223)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 7. Blood lymphatic cancers risk, voter record.

	Dependent variable:
	Deceased
Has social media account	1.066
	(0.955, 1.189)
Male	1.552
	(1.403, 1.716)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 8. All other cancers risk, voter record.

	Dependent variable:
	Deceased
Has social media account	1.040
	(0.996, 1.085)
Male	0.919
	(0.884, 0.956)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 9. Diabetes risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.624
	(0.557, 0.700)
Male	1.685
	(1.537, 1.847)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 10. Mental illness dementia risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.748
	(0.672, 0.833)
Male	1.437
	(1.313, 1.572)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 11. Ischemic heart disease risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.805
	(0.758, 0.855)
Male	3.019
	(2.853, 3.195)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 12. Stroke risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.713
	(0.632, 0.803)
Male	1.274
	(1.154, 1.406)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 13. all other CVDs risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.877
	(0.818, 0.940)
Male	1.984
	(1.864, 2.112)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 14. Liver disease risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.652
	(0.592, 0.718)
Male	1.847
	(1.707, 2.000)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 15. Unintentional injury risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.952
	(0.872, 1.040)
Male	3.190
	(2.918, 3.488)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 16. Drug overdose risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.987
	(0.900, 1.083)
Male	1.642
	(1.509, 1.787)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 17. Suicide risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.967
	(0.887, 1.054)
Male	3.089
	(2.831, 3.371)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 18. Homicide risk, voter record.

	Dependent variable:
	Deceased
Has social media account	0.554
	(0.457, 0.671)
Male	3.343
	(2.819, 3.965)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

F. Models included in Figure 1

Controls here included relationship status (baseline category: “not listed”) and application type used to access Facebook (baseline category: “desktop browser” or “no activity”).

Table 19. Friendships initiated

	Dependent variable:
	Deceased
Confirmation date	1.174 (1.150, 1.200)
Single	1.571 (1.487, 1.658)
In a relationship	1.165 (1.073, 1.265)
In an open relationship	1.793 (1.106, 2.908)
Married	0.758 (0.723, 0.794)
Engaged	1.031 (0.838, 1.269)
It’s complicated	1.959 (1.689, 2.271)
Widowed	1.941 (1.607, 2.344)
Separated	1.206 (0.751, 1.935)
Divorced	1.459 (1.046, 2.036)
Mobile browser	1.156 (1.013, 1.320)
Android	1.127 (0.974, 1.304)
iOS	1.034 (0.920, 1.163)
Blackberry	1.157 (0.968, 1.383)
Any smartphone	0.737 (0.695, 0.782)
Friendships initiated 0 [0-2]	1.156 (1.065, 1.255)
Friendships initiated 3 [2-6]	1.050 (0.972, 1.134)
Friendships initiated 7 [6-10]	1.041 (0.962, 1.127)
Friendships initiated 11 [10-14]	1.049 (0.967, 1.139)
Friendships initiated 24 [20-29]	0.995 (0.919, 1.078)
Friendships initiated 35 [29-43]	0.970 (0.895, 1.052)
Friendships initiated 53 [43-67]	1.023 (0.942, 1.110)
Friendships initiated 87 [67-122]	1.049 (0.965, 1.141)
Friendships initiated 195 [122-3414]	1.104 (1.013, 1.204)
Observations	89,597
Note:	95% confidence intervals estimated using robust standard errors.

Table 20. Friendships accepted

	Dependent variable:
	Deceased
Confirmation date	1.118 (1.093, 1.143)
Single	1.659 (1.572, 1.751)
In a relationship	1.242 (1.143, 1.349)
In an open relationship	1.844 (1.142, 2.978)
Married	0.804 (0.768, 0.842)
Engaged	1.095 (0.890, 1.347)
It's complicated	2.044 (1.763, 2.370)
Widowed	2.062 (1.708, 2.489)
Separated	1.269 (0.790, 2.037)
Divorced	1.485 (1.065, 2.071)
Mobile browser	1.169 (1.024, 1.335)
Android	1.101 (0.952, 1.274)
iOS	1.014 (0.902, 1.140)
Blackberry	1.142 (0.956, 1.365)
Any smartphone	0.778 (0.734, 0.825)
Friendships accepted 1 [0-4]	1.383 (1.280, 1.493)
Friendships accepted 6 [4-10]	1.211 (1.123, 1.307)
Friendships accepted 12 [10-16]	1.146 (1.061, 1.237)
Friendships accepted 18 [16-22]	1.091 (1.007, 1.182)
Friendships accepted 35 [31-41]	1.028 (0.948, 1.114)
Friendships accepted 47 [41-56]	0.962 (0.888, 1.043)
Friendships accepted 65 [56-79]	0.904 (0.831, 0.983)
Friendships accepted 98 [79-129]	0.852 (0.781, 0.929)
Friendships accepted 194 [129-4812]	0.915 (0.834, 1.003)
Observations	89,597
Note:	95% confidence intervals estimated using robust standard errors.

G. Models included in Figure 2

In models with two usage variables (heat maps shown in Figure 2), we interacted all deciles and included them as indicator variables. To present similar relative hazard scales (where a hazard of 1 corresponds to an average risk), we used the [0,0] interaction quantile as the reference category in the undirected analysis (no photos, no statuses) and the [1,1] interaction quantile as the reference category in the directed analysis (one sent post or message, one received photo tag).

Table 21. Statuses and photos (two-way decile estimates)

	Dependent variable:
	Deceased
Confirmation date	1.170
	(1.146, 1.195)
Single	1.522
	(1.441, 1.607)
In a relationship	1.153
	(1.062, 1.252)
In an open relationship	1.787
	(1.105, 2.889)
Married	0.745
	(0.712, 0.780)
Engaged	1.022
	(0.831, 1.257)
It's complicated	1.892
	(1.632, 2.194)
Widowed	1.878
	(1.554, 2.269)
Separated	1.198
	(0.746, 1.925)
Divorced	1.451
	(1.041, 2.024)
Mobile browser	1.073
	(0.939, 1.225)
Android	1.085
	(0.938, 1.254)
iOS	1.006
	(0.894, 1.131)
Blackberry	1.091
	(0.912, 1.304)
Any smartphone	0.768
	(0.723, 0.817)
Statuses/photos 0 1	0.803
	(0.735, 0.878)
Statuses/photos 0 11	0.702
	(0.601, 0.820)
Statuses/photos 0 124	0.673
	(0.519, 0.873)
Statuses/photos 0 33	0.599
	(0.478, 0.749)
Statuses/photos 0 4	0.751
	(0.659, 0.856)
Statuses/photos 1 0	1.011
	(0.927, 1.102)
Statuses/photos 1 1	0.983
	(0.855, 1.131)
Statuses/photos 1 11	0.753
	(0.615, 0.920)
Statuses/photos 1 124	0.549
	(0.386, 0.781)
Statuses/photos 1 33	0.672
	(0.519, 0.870)
Statuses/photos 1 4	0.938
	(0.791, 1.114)
Statuses/photos 20 0	1.283
	(1.095, 1.504)
Statuses/photos 20 1	1.040
	(0.869, 1.245)
Statuses/photos 20 11	0.946
	(0.834, 1.073)
Statuses/photos 20 124	0.771
	(0.670, 0.886)
Statuses/photos 20 33	0.894
	(0.788, 1.014)
Statuses/photos 20 4	1.083
	(0.925, 1.268)
Statuses/photos 3 0	1.068
	(0.972, 1.173)
Statuses/photos 3 1	1.001
	(0.878, 1.140)
Statuses/photos 3 11	0.835
	(0.719, 0.969)
Statuses/photos 3 124	0.674
	(0.531, 0.856)
Statuses/photos 3 33	0.711
	(0.588, 0.859)
Statuses/photos 3 4	0.927
	(0.802, 1.072)
Statuses/photos 7 0	1.147
	(1.009, 1.304)
Statuses/photos 7 1	1.019
	(0.874, 1.189)
Statuses/photos 7 11	0.973
	(0.852, 1.110)
Statuses/photos 7 124	0.788
	(0.656, 0.946)
Statuses/photos 7 33	0.892
	(0.769, 1.035)
Statuses/photos 7 4	0.972
	(0.834, 1.132)
Statuses/photos 72 0	1.630
	(1.276, 2.081)
Statuses/photos 72 1	1.485
	(1.174, 1.879)
Statuses/photos 72 11	1.195
	(1.045, 1.366)
Statuses/photos 72 124	0.959
	(0.866, 1.061)
Statuses/photos 72 33	1.060
	(0.944, 1.190)
Statuses/photos 72 4	1.154
	(0.938, 1.419)
Observations	89,597
Note:	95% confidence intervals estimated using robust standard errors.

Table 22. Posts/messages sent and photo tags received (two-way decile estimates)

	Dependent variable:
	Deceased
Confirmation date	1.162
	(1.138, 1.186)
Single	1.604
	(1.519, 1.694)
In a relationship	1.209
	(1.113, 1.313)
In an open relationship	1.838
	(1.137, 2.971)
Married	0.783
	(0.748, 0.820)
Engaged	1.078
	(0.876, 1.327)
It's complicated	2.005
	(1.730, 2.325)
Widowed	2.015
	(1.669, 2.433)
Separated	1.204
	(0.749, 1.936)
Divorced	1.497
	(1.075, 2.085)
Mobile browser	1.163
	(1.019, 1.328)
Android	1.127
	(0.973, 1.305)
iOS	1.030
	(0.915, 1.159)
Blackberry	1.146
	(0.959, 1.370)
Any smartphone	0.775
	(0.730, 0.823)
Posts messages out/tags in 0 0	1.147
	(1.012, 1.301)
Posts messages out/tags in 0 1	1.017
	(0.882, 1.173)
Posts messages out/tags in 0 14	0.668
	(0.426, 1.046)
Posts messages out/tags in 0 2	0.941
	(0.808, 1.096)
Posts messages out/tags in 0 4	0.810
	(0.649, 1.010)
Posts messages out/tags in 0 7	0.870
	(0.669, 1.130)
Posts messages out/tags in 10 0	1.095
	(0.935, 1.283)
Posts messages out/tags in 10 1	0.905
	(0.760, 1.078)
Posts messages out/tags in 10 14	0.738
	(0.548, 0.992)
Posts messages out/tags in 10 2	0.843
	(0.711, 0.998)
Posts messages out/tags in 10 4	0.825
	(0.667, 1.019)
Posts messages out/tags in 10 7	0.806
	(0.641, 1.014)
Posts messages out/tags in 150 0	0.965
	(0.663, 1.404)
Posts messages out/tags in 150 1	1.140
	(0.875, 1.484)
Posts messages out/tags in 150 14	0.866
	(0.741, 1.011)
Posts messages out/tags in 150 2	1.038
	(0.855, 1.261)
Posts messages out/tags in 150 4	0.941
	(0.774, 1.145)
Posts messages out/tags in 150 7	0.989
	(0.833, 1.174)
Continued	
Posts messages out/tags in 18 0	0.992
	(0.837, 1.176)
Posts messages out/tags in 18 1	1.012
	(0.851, 1.203)
Posts messages out/tags in 18 14	0.838
	(0.664, 1.058)
Posts messages out/tags in 18 2	0.896
	(0.759, 1.058)
Posts messages out/tags in 18 4	0.794
	(0.648, 0.972)
Posts messages out/tags in 18 7	0.802
	(0.651, 0.988)
Posts messages out/tags in 2 0	1.065
	(0.920, 1.232)
Posts messages out/tags in 2 14	0.740
	(0.464, 1.180)
Posts messages out/tags in 2 2	0.811
	(0.675, 0.973)
Posts messages out/tags in 2 4	0.917
	(0.711, 1.183)
Posts messages out/tags in 2 7	0.789
	(0.579, 1.074)
Posts messages out/tags in 33 0	0.943
	(0.778, 1.143)
Posts messages out/tags in 33 1	1.020
	(0.848, 1.226)
Posts messages out/tags in 33 14	0.802
	(0.651, 0.989)
Posts messages out/tags in 33 2	1.043
	(0.887, 1.227)
Posts messages out/tags in 33 4	0.825
	(0.682, 0.999)
Posts messages out/tags in 33 7	0.815
	(0.669, 0.992)
Posts messages out/tags in 5 0	1.011
	(0.871, 1.174)
Posts messages out/tags in 5 1	0.950
	(0.807, 1.119)
Posts messages out/tags in 5 14	0.985
	(0.723, 1.343)
Posts messages out/tags in 5 2	0.751
	(0.632, 0.893)
Posts messages out/tags in 5 4	0.783
	(0.626, 0.979)
Posts messages out/tags in 5 7	0.773
	(0.601, 0.993)
Posts messages out/tags in 62 0	1.008
	(0.803, 1.265)
Posts messages out/tags in 62 1	0.853
	(0.687, 1.059)
Posts messages out/tags in 62 14	0.854
	(0.715, 1.021)
Posts messages out/tags in 62 2	1.060
	(0.897, 1.253)
Posts messages out/tags in 62 4	0.864
	(0.717, 1.041)
Posts messages out/tags in 62 7	0.770
	(0.639, 0.928)
Observations	89,597
Note:	95% confidence intervals estimated using robust standard errors.

H. Continuous estimates for relationships in Figure 2

Table 23. Statuses and photos (same model)

	Dependent variable:
	Deceased
Confirmation date	1.170
	(1.146, 1.195)
Single	1.507
	(1.427, 1.591)
In a relationship	1.143
	(1.053, 1.241)
In an open relationship	1.772
	(1.094, 2.870)
Married	0.739
	(0.706, 0.773)
Engaged	1.014
	(0.824, 1.248)
It's complicated	1.882
	(1.623, 2.181)
Widowed	1.851
	(1.532, 2.236)
Separated	1.186
	(0.738, 1.904)
Divorced	1.427
	(1.024, 1.989)
Mobile browser	1.058
	(0.926, 1.208)
Android	1.067
	(0.923, 1.234)
iOS	0.997
	(0.886, 1.121)
Blackberry	1.078
	(0.901, 1.288)
Any smartphone	0.766
	(0.721, 0.814)
Statuses	1.167
	(1.137, 1.197)
Photos	0.836
	(0.814, 0.858)
Observations	89,597

Note: 95% confidence intervals estimated using robust standard errors.

Table 24. Posts/messages sent (non-linear estimate)

	Dependent variable:
	Deceased
Confirmation date	1.167
	(1.143, 1.191)
Single	1.599
	(1.514, 1.688)
In a relationship	1.189
	(1.095, 1.291)
In an open relationship	1.815
	(1.122, 2.936)
Married	0.767
	(0.733, 0.803)
Engaged	1.048
	(0.852, 1.290)
It's complicated	1.986
	(1.713, 2.303)
Widowed	1.988
	(1.647, 2.400)
Separated	1.222
	(0.761, 1.962)
Divorced	1.488
	(1.066, 2.076)
Mobile browser	1.166
	(1.021, 1.331)
Android	1.107
	(0.957, 1.282)
iOS	1.024
	(0.910, 1.152)
Blackberry	1.152
	(0.964, 1.377)
Any smartphone	0.762
	(0.718, 0.809)
Posts/messages sent	0.804
	(0.761, 0.849)
Posts/messages sent^2	1.196
	(1.133, 1.263)
Observations	89,597
Note:	95% confidence intervals estimated using robust standard errors.

Table 25. Photo tags received (non-linear estimate)

	Dependent variable:
	Deceased
Confirmation date	1.160
	(1.136, 1.184)
Single	1.598
	(1.515, 1.685)
In a relationship	1.201
	(1.107, 1.304)
In an open relationship	1.843
	(1.140, 2.979)
Married	0.776
	(0.742, 0.812)
Engaged	1.066
	(0.867, 1.311)
It's complicated	1.992
	(1.718, 2.309)
Widowed	2.001
	(1.658, 2.416)
Separated	1.214
	(0.756, 1.951)
Divorced	1.479
	(1.062, 2.059)
Mobile browser	1.161
	(1.017, 1.325)
Android	1.100
	(0.951, 1.272)
iOS	1.005
	(0.894, 1.130)
Blackberry	1.134
	(0.949, 1.355)
Any smartphone	0.778
	(0.734, 0.825)
Tags received	0.830
	(0.790, 0.872)
Tags received^2	1.100
	(1.046, 1.157)
Observations	89,597
Note:	95% confidence intervals estimated using robust standard errors.

I. Models included in Figure 3

Table 26. Facebook friend count

	Dependent variable:			
	Cancer	CVD	Drug overdose	Suicide
	(1)	(2)	(3)	(4)
Confirmation date	1.045 (1.008, 1.083)	1.220 (1.159, 1.285)	1.249 (1.143, 1.365)	0.981 (0.911, 1.057)
Single	1.205 (1.094, 1.327)	1.884 (1.678, 2.114)	2.418 (1.921, 3.043)	1.790 (1.414, 2.267)
In a relationship	1.003 (0.861, 1.169)	1.187 (0.979, 1.441)	2.242 (1.680, 2.993)	1.262 (0.920, 1.733)
In an open relationship	1.270 (0.471, 3.423)	1.483 (0.472, 4.658)		1.680 (0.241, 11.730)
Married	0.899 (0.838, 0.965)	0.719 (0.648, 0.797)	0.673 (0.522, 0.867)	0.709 (0.569, 0.883)
Engaged	0.741 (0.464, 1.181)	1.239 (0.782, 1.965)	1.381 (0.701, 2.719)	1.181 (0.601, 2.322)
It's complicated	1.689 (1.306, 2.184)	2.061 (1.493, 2.845)	4.181 (2.564, 6.817)	2.175 (1.156, 4.091)
Widowed	1.397 (1.009, 1.934)	3.189 (2.284, 4.454)	4.103 (1.683, 10.005)	2.087 (0.666, 6.543)
Separated	1.193 (0.537, 2.647)	1.040 (0.335, 3.227)	1.501 (0.210, 10.711)	1.321 (0.183, 9.533)
Divorced	0.732 (0.349, 1.536)	0.766 (0.288, 2.036)	2.164 (0.532, 8.803)	4.654 (1.916, 11.306)
Mobile browser	1.035 (0.804, 1.333)	1.249 (0.931, 1.677)	0.680 (0.382, 1.209)	1.094 (0.664, 1.803)
Android	1.290 (1.015, 1.638)	0.843 (0.578, 1.229)	1.071 (0.612, 1.873)	1.139 (0.708, 1.833)
iOS	1.164 (0.972, 1.394)	0.939 (0.716, 1.232)	1.173 (0.726, 1.895)	0.837 (0.533, 1.314)
Blackberry	1.176 (0.881, 1.569)	1.132 (0.763, 1.680)	1.121 (0.511, 2.456)	0.540 (0.221, 1.320)
Any smartphone	0.869 (0.791, 0.954)	0.803 (0.705, 0.915)	0.512 (0.394, 0.665)	0.903 (0.720, 1.132)
Friend count	0.991 (0.955, 1.029)	0.910 (0.866, 0.957)	0.779 (0.701, 0.866)	0.731 (0.663, 0.805)
Observations	89,597	89,597	89,597	89,597
Note:	95% confidence intervals estimated using robust standard errors.			

Table 27. Friendships initiated

	Dependent variable:			
	Cancer	CVD	Drug overdose	Suicide
	(1)	(2)	(3)	(4)
Confirmation date	1.050 (1.014, 1.088)	1.254 (1.193, 1.317)	1.313 (1.205, 1.430)	1.036 (0.964, 1.114)
Single	1.188 (1.078, 1.309)	1.795 (1.600, 2.015)	2.233 (1.774, 2.810)	1.661 (1.314, 2.100)
In a relationship	0.988 (0.848, 1.151)	1.124 (0.927, 1.364)	2.045 (1.536, 2.723)	1.155 (0.846, 1.578)
In an open relationship	1.251 (0.464, 3.374)	1.431 (0.456, 4.494)		1.585 (0.224, 11.211)
Married	0.887 (0.826, 0.951)	0.683 (0.616, 0.756)	0.616 (0.480, 0.792)	0.648 (0.523, 0.803)
Engaged	0.730 (0.457, 1.164)	1.174 (0.740, 1.861)	1.255 (0.637, 2.472)	1.079 (0.550, 2.119)
It's complicated	1.666 (1.288, 2.155)	1.971 (1.428, 2.721)	3.940 (2.418, 6.420)	2.050 (1.091, 3.850)
Widowed	1.373 (0.991, 1.901)	3.016 (2.159, 4.213)	3.769 (1.546, 9.191)	1.925 (0.614, 6.038)
Separated	1.173 (0.528, 2.606)	0.987 (0.318, 3.066)	1.383 (0.193, 9.887)	1.222 (0.169, 8.833)
Divorced	0.723 (0.345, 1.518)	0.742 (0.279, 1.972)	2.087 (0.514, 8.480)	4.546 (1.872, 11.042)
Mobile browser	1.034 (0.803, 1.332)	1.239 (0.923, 1.663)	0.667 (0.375, 1.186)	1.075 (0.653, 1.770)
Android	1.300 (1.023, 1.651)	0.861 (0.590, 1.256)	1.091 (0.624, 1.907)	1.158 (0.719, 1.864)
iOS	1.172 (0.979, 1.404)	0.956 (0.729, 1.255)	1.192 (0.737, 1.929)	0.845 (0.538, 1.328)
Blackberry	1.182 (0.886, 1.578)	1.145 (0.771, 1.700)	1.128 (0.514, 2.472)	0.542 (0.222, 1.326)
Any smartphone	0.858 (0.781, 0.942)	0.769 (0.675, 0.877)	0.480 (0.370, 0.623)	0.846 (0.677, 1.057)
Friendships initiated	1.016 (0.981, 1.053)	0.990 (0.944, 1.039)	0.894 (0.805, 0.993)	0.840 (0.764, 0.923)
Observations	89,597	89,597	89,597	89,597

Note: 95% confidence intervals estimated using robust standard errors.

Table 28. Friendships accepted

	Dependent variable:			
	Cancer	CVD	Drug overdose	Suicide
	(1)	(2)	(3)	(4)
Confirmation date	1.042	1.188	1.203	0.948
	(1.004, 1.082)	(1.128, 1.252)	(1.100, 1.315)	(0.879, 1.022)
Single	1.207	1.910	2.452	1.790
	(1.096, 1.328)	(1.703, 2.142)	(1.953, 3.079)	(1.415, 2.264)
In a relationship	1.005	1.210	2.302	1.278
	(0.863, 1.170)	(0.997, 1.467)	(1.728, 3.066)	(0.931, 1.754)
In an open relationship	1.270	1.480		1.657
	(0.471, 3.424)	(0.471, 4.651)		(0.243, 11.325)
Married	0.901	0.732	0.688	0.716
	(0.840, 0.966)	(0.661, 0.810)	(0.535, 0.886)	(0.574, 0.892)
Engaged	0.742	1.258	1.409	1.192
	(0.465, 1.183)	(0.794, 1.994)	(0.717, 2.770)	(0.607, 2.341)
It's complicated	1.691	2.078	4.187	2.160
	(1.308, 2.186)	(1.506, 2.868)	(2.564, 6.837)	(1.148, 4.064)
Widowed	1.399	3.229	4.104	2.053
	(1.010, 1.937)	(2.314, 4.507)	(1.685, 9.997)	(0.655, 6.435)
Separated	1.194	1.051	1.525	1.315
	(0.538, 2.650)	(0.339, 3.259)	(0.214, 10.875)	(0.182, 9.485)
Divorced	0.732	0.767	2.157	4.517
	(0.349, 1.537)	(0.289, 2.037)	(0.530, 8.778)	(1.863, 10.951)
Mobile browser	1.036	1.257	0.685	1.101
	(0.804, 1.334)	(0.936, 1.687)	(0.385, 1.217)	(0.668, 1.814)
Android	1.289	0.843	1.082	1.158
	(1.015, 1.637)	(0.578, 1.230)	(0.618, 1.893)	(0.720, 1.864)
iOS	1.164	0.940	1.186	0.853
	(0.972, 1.393)	(0.716, 1.232)	(0.735, 1.914)	(0.544, 1.338)
Blackberry	1.176	1.137	1.130	0.547
	(0.881, 1.569)	(0.766, 1.687)	(0.515, 2.477)	(0.224, 1.337)
Any smartphone	0.870	0.813	0.520	0.906
	(0.793, 0.955)	(0.714, 0.925)	(0.401, 0.675)	(0.722, 1.136)
Friendships accepted	0.986	0.869	0.733	0.696
	(0.951, 1.023)	(0.829, 0.912)	(0.664, 0.808)	(0.634, 0.765)
Observations	89,597	89,597	89,597	89,597

Note: 95% confidence intervals estimated using robust standard errors.

Table 29. Statuses

	Dependent variable:			
	Cancer	CVD	Drug overdose	Suicide
	(1)	(2)	(3)	(4)
Confirmation date	1.054	1.263	1.336	1.069
	(1.018, 1.091)	(1.202, 1.326)	(1.228, 1.453)	(0.994, 1.149)
Single	1.159	1.722	2.145	1.449
	(1.052, 1.277)	(1.533, 1.934)	(1.704, 2.701)	(1.146, 1.831)
In a relationship	0.968	1.083	1.954	1.017
	(0.832, 1.127)	(0.894, 1.313)	(1.477, 2.586)	(0.748, 1.381)
In an open relationship	1.232	1.392		1.481
	(0.457, 3.317)	(0.444, 4.363)		(0.209, 10.476)
Married	0.872	0.663	0.588	0.574
	(0.814, 0.935)	(0.600, 0.733)	(0.460, 0.750)	(0.465, 0.708)
Engaged	0.717	1.133	1.189	0.950
	(0.450, 1.143)	(0.716, 1.795)	(0.606, 2.334)	(0.484, 1.866)
It's complicated	1.627	1.895	3.824	1.796
	(1.259, 2.104)	(1.373, 2.616)	(2.343, 6.240)	(0.952, 3.389)
Widowed	1.334	2.873	3.601	1.630
	(0.964, 1.847)	(2.057, 4.014)	(1.475, 8.792)	(0.519, 5.116)
Separated	1.154	0.953	1.311	1.067
	(0.520, 2.560)	(0.307, 2.961)	(0.183, 9.376)	(0.148, 7.713)
Divorced	0.703	0.713	2.043	4.121
	(0.335, 1.475)	(0.268, 1.895)	(0.504, 8.292)	(1.699, 9.993)
Mobile browser	1.020	1.217	0.675	1.035
	(0.792, 1.314)	(0.906, 1.635)	(0.378, 1.204)	(0.627, 1.708)
Android	1.292	0.865	1.129	1.216
	(1.018, 1.641)	(0.593, 1.260)	(0.645, 1.974)	(0.756, 1.957)
iOS	1.186	0.976	1.217	0.904
	(0.990, 1.421)	(0.744, 1.281)	(0.752, 1.968)	(0.576, 1.417)
Blackberry	1.182	1.155	1.154	0.569
	(0.885, 1.577)	(0.778, 1.714)	(0.527, 2.527)	(0.233, 1.392)
Any smartphone	0.826	0.727	0.470	0.726
	(0.751, 0.909)	(0.637, 0.830)	(0.362, 0.611)	(0.578, 0.913)
Statuses	1.064	1.068	0.956	1.066
	(1.028, 1.101)	(1.018, 1.121)	(0.875, 1.045)	(0.972, 1.169)
Observations	89,597	89,597	89,597	89,597

Note: 95% confidence intervals estimated using robust standard errors.

Table 30. Photos

	Dependent variable:			
	Cancer	CVD	Drug overdose	Suicide
	(1)	(2)	(3)	(4)
Confirmation date	1.043	1.246	1.326	1.065
	(1.008, 1.080)	(1.186, 1.308)	(1.219, 1.443)	(0.990, 1.145)
Single	1.218	1.853	2.228	1.522
	(1.108, 1.340)	(1.655, 2.075)	(1.779, 2.792)	(1.210, 1.913)
In a relationship	1.016	1.169	2.055	1.058
	(0.873, 1.183)	(0.964, 1.417)	(1.554, 2.718)	(0.779, 1.436)
In an open relationship	1.286	1.471		1.529
	(0.477, 3.468)	(0.468, 4.626)		(0.216, 10.817)
Married	0.909	0.701	0.612	0.592
	(0.849, 0.973)	(0.635, 0.774)	(0.480, 0.781)	(0.480, 0.730)
Engaged	0.751	1.219	1.249	0.987
	(0.471, 1.196)	(0.769, 1.930)	(0.636, 2.452)	(0.504, 1.932)
It's complicated	1.710	2.033	3.972	1.888
	(1.322, 2.210)	(1.474, 2.805)	(2.444, 6.454)	(1.003, 3.554)
Widowed	1.415	3.137	3.777	1.731
	(1.023, 1.959)	(2.248, 4.377)	(1.549, 9.207)	(0.554, 5.413)
Separated	1.208	1.024	1.400	1.109
	(0.544, 2.682)	(0.330, 3.175)	(0.196, 9.987)	(0.154, 7.989)
Divorced	0.745	0.777	2.140	4.301
	(0.355, 1.564)	(0.293, 2.065)	(0.526, 8.701)	(1.776, 10.420)
Mobile browser	1.023	1.210	0.649	1.058
	(0.794, 1.318)	(0.901, 1.624)	(0.365, 1.153)	(0.642, 1.742)
Android	1.281	0.846	1.080	1.213
	(1.008, 1.626)	(0.580, 1.233)	(0.617, 1.893)	(0.753, 1.955)
iOS	1.145	0.919	1.141	0.878
	(0.955, 1.371)	(0.700, 1.206)	(0.704, 1.850)	(0.560, 1.377)
Blackberry	1.155	1.092	1.082	0.559
	(0.865, 1.542)	(0.735, 1.623)	(0.493, 2.377)	(0.229, 1.367)
Any smartphone	0.902	0.848	0.541	0.789
	(0.820, 0.993)	(0.740, 0.971)	(0.414, 0.707)	(0.626, 0.993)
Photos	0.954	0.893	0.821	0.968
	(0.921, 0.987)	(0.849, 0.938)	(0.748, 0.901)	(0.885, 1.059)
Observations	89,597	89,597	89,597	89,597

Note: 95% confidence intervals estimated using robust standard errors.

Table 31. Posts/messages sent

	Dependent variable:			
	Cancer	CVD	Drug overdose	Suicide
	(1)	(2)	(3)	(4)
Confirmation date	1.044	1.251	1.324	1.054
	(1.008, 1.081)	(1.191, 1.314)	(1.215, 1.441)	(0.979, 1.134)
Single	1.214	1.816	2.204	1.596
	(1.103, 1.338)	(1.617, 2.039)	(1.755, 2.769)	(1.264, 2.016)
In a relationship	1.010	1.136	2.004	1.095
	(0.867, 1.176)	(0.936, 1.378)	(1.514, 2.652)	(0.806, 1.488)
In an open relationship	1.276	1.441		1.569
	(0.473, 3.442)	(0.459, 4.526)		(0.222, 11.082)
Married	0.904	0.687	0.599	0.607
	(0.843, 0.968)	(0.621, 0.759)	(0.470, 0.763)	(0.493, 0.748)
Engaged	0.745	1.183	1.217	1.019
	(0.467, 1.187)	(0.747, 1.873)	(0.620, 2.389)	(0.520, 1.996)
It's complicated	1.700	1.988	3.881	1.959
	(1.314, 2.199)	(1.440, 2.744)	(2.388, 6.306)	(1.041, 3.686)
Widowed	1.408	3.049	3.704	1.824
	(1.017, 1.949)	(2.183, 4.257)	(1.517, 9.046)	(0.583, 5.709)
Separated	1.200	0.997	1.338	1.153
	(0.541, 2.664)	(0.321, 3.092)	(0.187, 9.573)	(0.160, 8.316)
Divorced	0.739	0.752	2.094	4.456
	(0.352, 1.552)	(0.283, 1.999)	(0.515, 8.506)	(1.837, 10.812)
Mobile browser	1.036	1.243	0.676	1.086
	(0.805, 1.334)	(0.926, 1.668)	(0.380, 1.204)	(0.660, 1.789)
Android	1.274	0.847	1.060	1.139
	(1.002, 1.620)	(0.581, 1.235)	(0.604, 1.861)	(0.704, 1.843)
iOS	1.152	0.944	1.171	0.840
	(0.961, 1.381)	(0.719, 1.239)	(0.721, 1.900)	(0.535, 1.319)
Blackberry	1.169	1.137	1.132	0.549
	(0.876, 1.561)	(0.766, 1.688)	(0.516, 2.482)	(0.224, 1.343)
Any smartphone	0.879	0.782	0.484	0.828
	(0.799, 0.967)	(0.685, 0.892)	(0.372, 0.630)	(0.659, 1.042)
Posts/messages sent	0.979	0.971	0.910	0.898
	(0.946, 1.013)	(0.926, 1.019)	(0.828, 0.999)	(0.815, 0.990)
Observations	89,597	89,597	89,597	89,597

Note: 95% confidence intervals estimated using robust standard errors.

Table 32. Posts/messages received

	Dependent variable:			
	Cancer	CVD	Drug overdose	Suicide
	(1)	(2)	(3)	(4)
Confirmation date	1.060	1.252	1.316	1.040
	(1.024, 1.098)	(1.192, 1.316)	(1.209, 1.433)	(0.967, 1.118)
Single	1.149	1.802	2.252	1.686
	(1.043, 1.265)	(1.604, 2.024)	(1.792, 2.830)	(1.333, 2.132)
In a relationship	0.954	1.128	2.063	1.167
	(0.819, 1.111)	(0.929, 1.370)	(1.554, 2.739)	(0.855, 1.592)
In an open relationship	1.224	1.433		1.596
	(0.455, 3.295)	(0.456, 4.499)		(0.228, 11.195)
Married	0.863	0.684	0.613	0.638
	(0.806, 0.925)	(0.618, 0.756)	(0.481, 0.783)	(0.516, 0.788)
Engaged	0.705	1.177	1.259	1.091
	(0.442, 1.124)	(0.743, 1.866)	(0.641, 2.475)	(0.556, 2.142)
It's complicated	1.614	1.977	3.990	2.077
	(1.248, 2.087)	(1.432, 2.729)	(2.454, 6.488)	(1.102, 3.914)
Widowed	1.325	3.024	3.796	1.940
	(0.957, 1.834)	(2.165, 4.225)	(1.554, 9.273)	(0.620, 6.071)
Separated	1.132	0.991	1.390	1.247
	(0.511, 2.509)	(0.319, 3.077)	(0.194, 9.946)	(0.173, 8.999)
Divorced	0.696	0.746	2.130	4.671
	(0.332, 1.461)	(0.281, 1.982)	(0.524, 8.665)	(1.928, 11.316)
Mobile browser	1.028	1.240	0.676	1.095
	(0.799, 1.324)	(0.924, 1.665)	(0.380, 1.201)	(0.665, 1.803)
Android	1.326	0.859	1.085	1.147
	(1.044, 1.684)	(0.589, 1.253)	(0.620, 1.900)	(0.712, 1.848)
iOS	1.195	0.954	1.183	0.836
	(0.998, 1.431)	(0.727, 1.252)	(0.731, 1.913)	(0.533, 1.311)
Blackberry	1.201	1.143	1.127	0.539
	(0.900, 1.603)	(0.771, 1.696)	(0.514, 2.471)	(0.220, 1.321)
Any smartphone	0.825	0.773	0.493	0.880
	(0.751, 0.907)	(0.678, 0.881)	(0.380, 0.639)	(0.701, 1.105)
Posts/messages received	1.078	0.985	0.877	0.815
	(1.041, 1.116)	(0.939, 1.032)	(0.798, 0.964)	(0.741, 0.897)
Observations	89,597	89,597	89,597	89,597

Note: 95% confidence intervals estimated using robust standard errors.

Table 33. Photo tags received

	Dependent variable:			
	Cancer	CVD	Drug overdose	Suicide
	(1)	(2)	(3)	(4)
Confirmation date	1.048	1.244	1.286	1.026
	(1.013, 1.086)	(1.184, 1.306)	(1.181, 1.399)	(0.955, 1.103)
Single	1.196	1.833	2.310	1.655
	(1.087, 1.315)	(1.637, 2.052)	(1.844, 2.894)	(1.317, 2.080)
In a relationship	0.994	1.163	2.239	1.215
	(0.854, 1.157)	(0.960, 1.410)	(1.687, 2.971)	(0.895, 1.649)
In an open relationship	1.259	1.472		1.637
	(0.467, 3.391)	(0.469, 4.626)		(0.235, 11.384)
Married	0.892	0.703	0.656	0.661
	(0.833, 0.955)	(0.636, 0.776)	(0.513, 0.840)	(0.536, 0.816)
Engaged	0.734	1.220	1.398	1.161
	(0.460, 1.170)	(0.770, 1.933)	(0.714, 2.740)	(0.591, 2.278)
It's complicated	1.676	2.018	4.112	2.073
	(1.297, 2.167)	(1.463, 2.784)	(2.523, 6.702)	(1.103, 3.898)
Widowed	1.383	3.110	3.982	1.934
	(0.999, 1.915)	(2.230, 4.339)	(1.634, 9.705)	(0.618, 6.048)
Separated	1.184	1.000	1.362	1.153
	(0.534, 2.626)	(0.322, 3.099)	(0.192, 9.655)	(0.160, 8.314)
Divorced	0.727	0.758	2.171	4.534
	(0.346, 1.526)	(0.286, 2.012)	(0.535, 8.817)	(1.877, 10.949)
Mobile browser	1.034	1.247	0.699	1.111
	(0.803, 1.332)	(0.929, 1.674)	(0.393, 1.242)	(0.675, 1.829)
Android	1.295	0.849	1.053	1.141
	(1.020, 1.645)	(0.582, 1.237)	(0.603, 1.841)	(0.709, 1.837)
iOS	1.169	0.938	1.106	0.806
	(0.976, 1.400)	(0.715, 1.231)	(0.686, 1.784)	(0.515, 1.264)
Blackberry	1.179	1.130	1.108	0.533
	(0.884, 1.574)	(0.762, 1.677)	(0.505, 2.428)	(0.218, 1.301)
Any smartphone	0.861	0.807	0.568	0.959
	(0.784, 0.945)	(0.709, 0.918)	(0.438, 0.736)	(0.767, 1.200)
Photo tags received	1.009	0.910	0.700	0.695
	(0.976, 1.044)	(0.868, 0.954)	(0.636, 0.770)	(0.632, 0.765)
Observations	89,597	89,597	89,597	89,597

Note: 95% confidence intervals estimated using robust standard errors.

J. Extended cause-specific estimates for models included in Figure 3

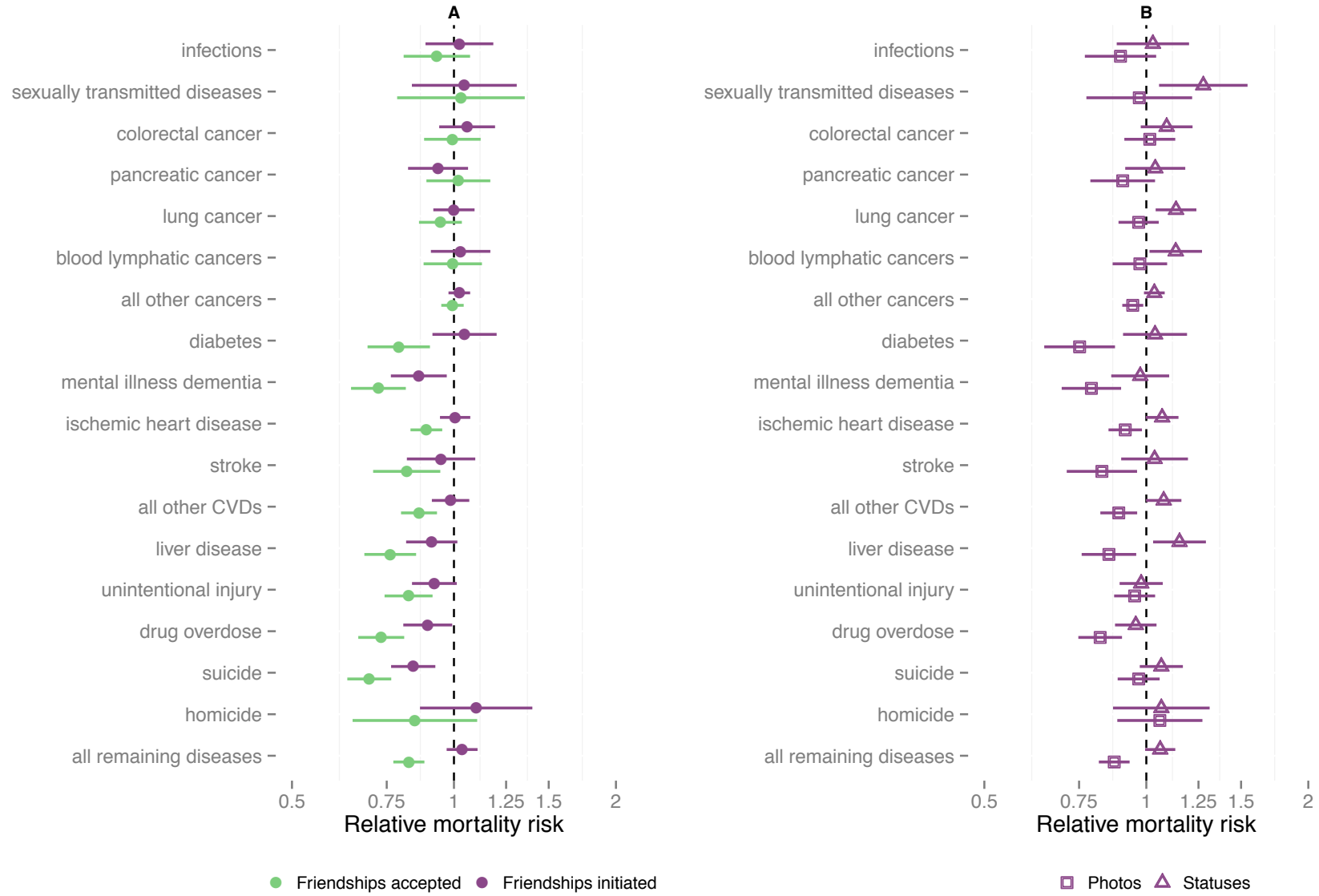

Fig. 8. *Facebook activities, by directionality and activity type (seventeen cause of death categories).*

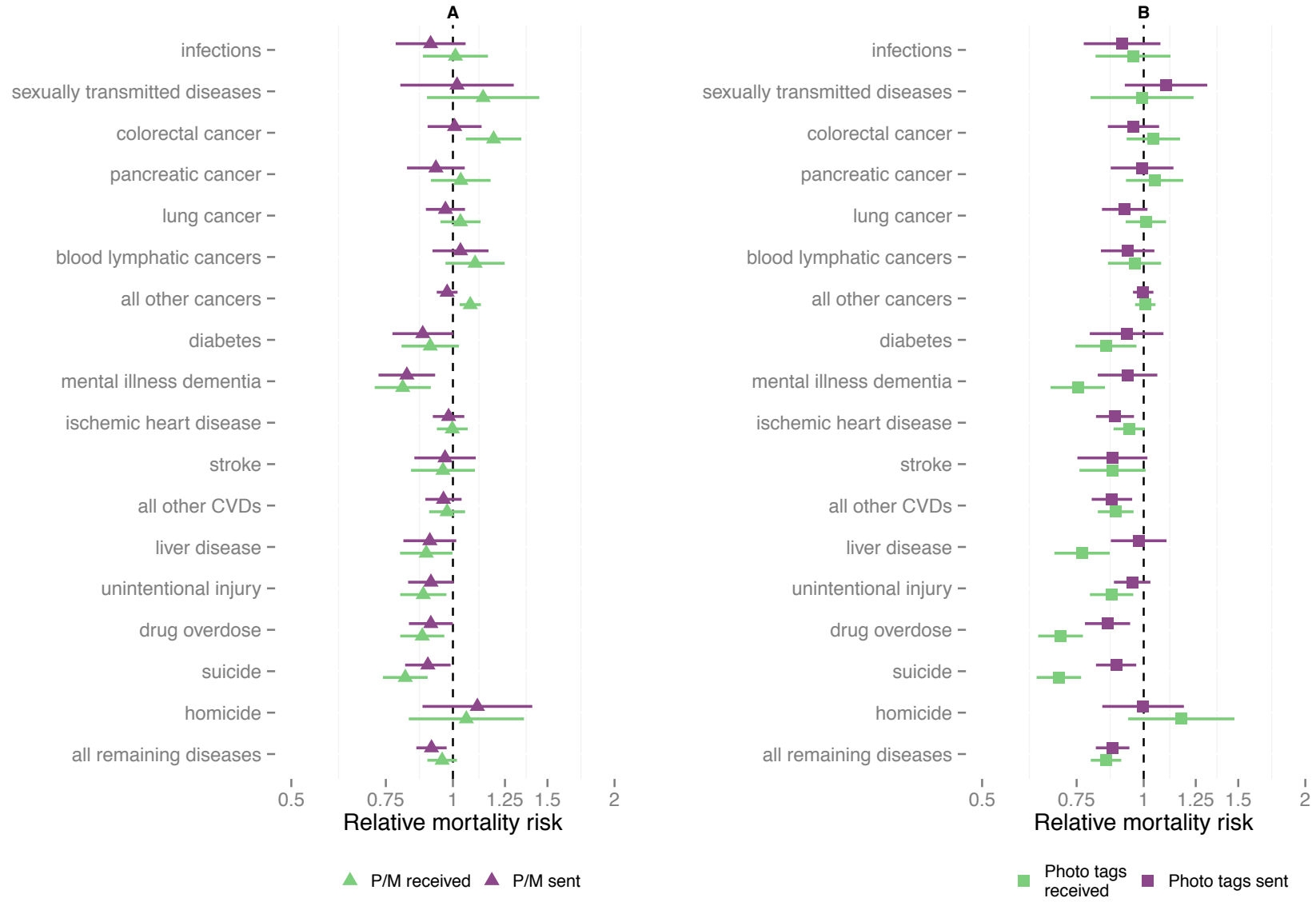

Fig. 9. *Facebook activities, by directionality and activity type (seventeen cause of death categories).*

K. Results for full (non-voter included) population

For these full sample replication models, we added control variables for the proportion of Facebook friends who, during the observation period, took any action on the site through a common, recognizable computer application (desktop browser, mobile browser, iOS app, Android app, Blackberry app), listed education (three variables for proportions of friends listing high school, college, and graduate school education), and an indicator variable for whether the individual was present in the California voter record.

Some profiles in the sample appeared to have many Facebook friends who were completely inactive. This friend inactivity is consistent with clusters of fake profiles. Only 10% of profiles in this Facebook sample had fewer than 90% of their friends take an action on the website over the six-month period. To more clearly compare these results to those in the main paper, we also display the 'voter' subsample with these additional controls.

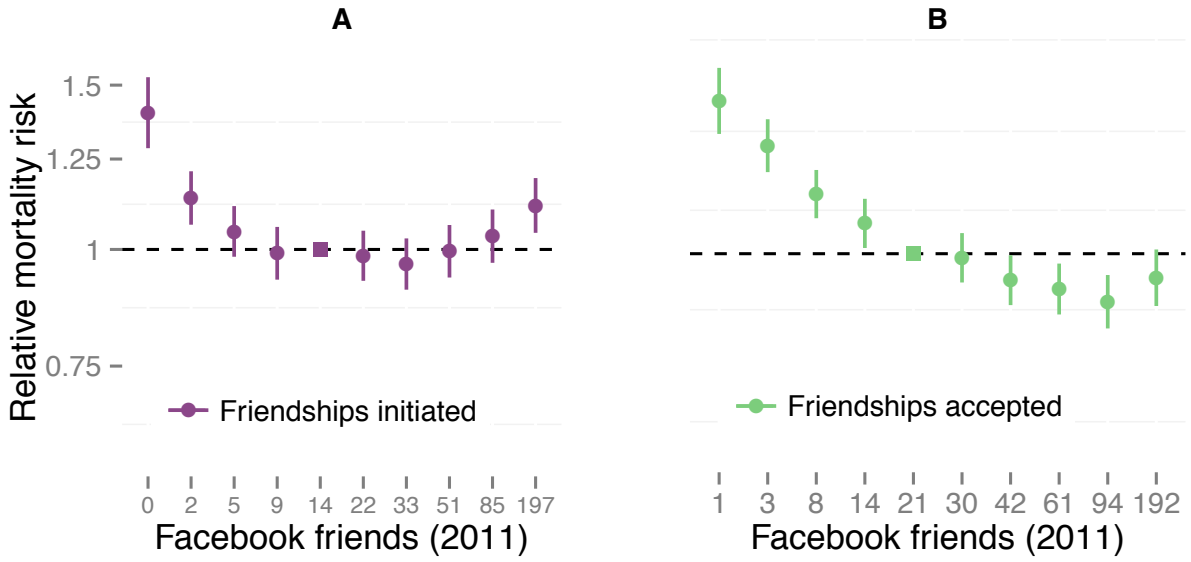

Fig. 10. *Facebook friends and relative mortality risk (all-cause mortality).* This figure shows all-cause mortality regressed on deciles of Facebook friend counts, by initiated and received Facebook friendships. These results replicate Figure 1 for the full, including non-voting population.

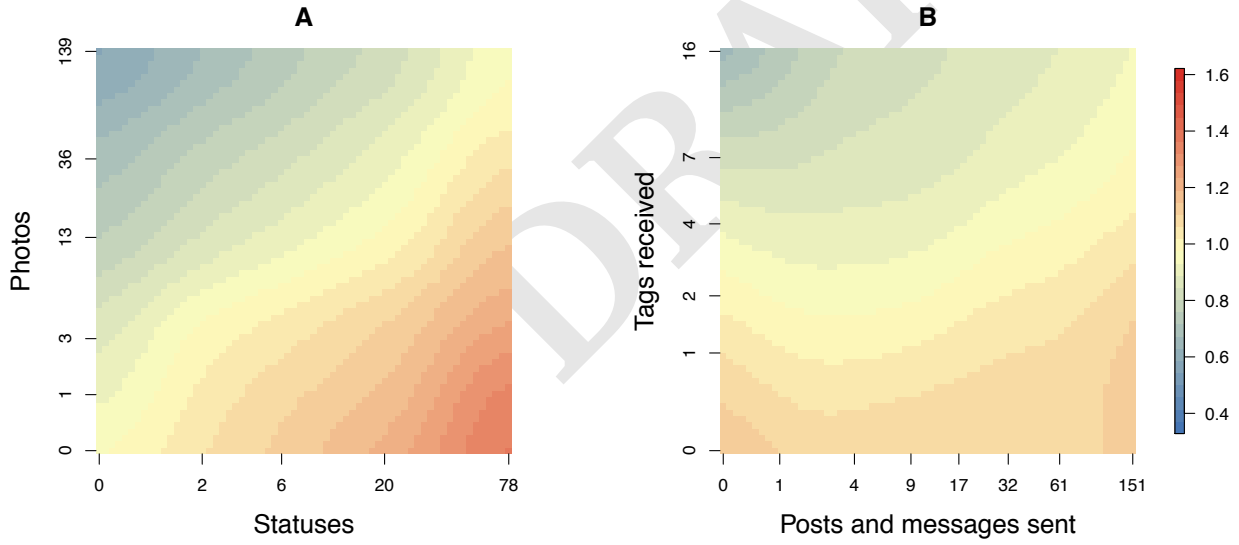

Fig. 11. *Facebook activities, smoothed mortality risk by quantile activity-activity interaction.* This figure shows all-cause mortality risk among Facebook users for combinations of activities. These results replicate Figure 2 for the full, including non-voting population. Controlling for attributes of Facebook friends (especially control variables related to friends' first and last name characteristics and device types) increases the non-linearity of the relationship for posts and messages sent.

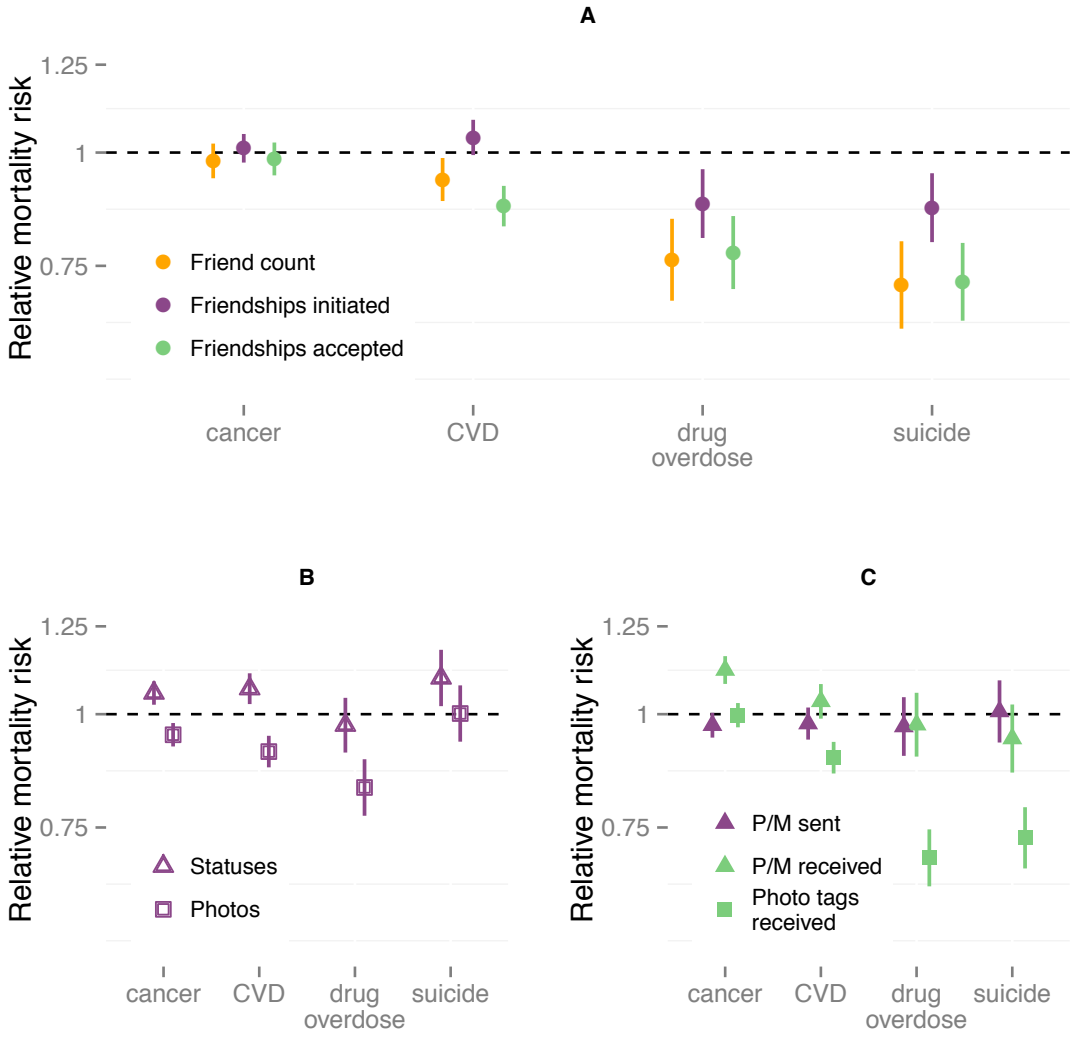

Fig. 12. *Facebook activities, by directionality and activity type.* This figure shows cause-specific morality risk among Facebook users by online activity levels (counts of Facebook tool use), separated by the directionality of the action (from or to the subject) and activity type (the type of Facebook tool used). These results replicate Figure 3 for the full, including non-voting population.

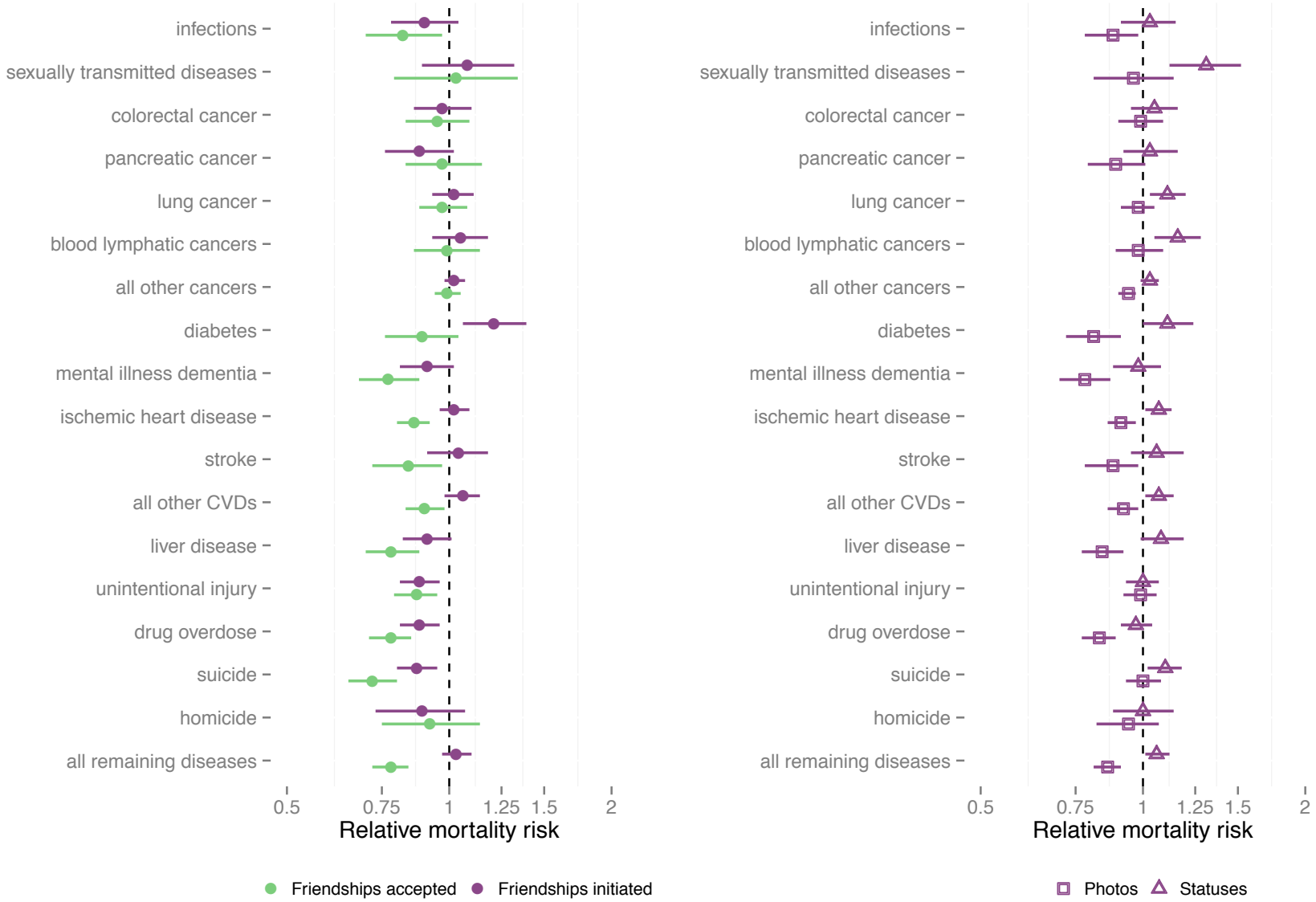

Fig. 13. *Facebook activities, by directionality and activity type (seventeen cause of death categories).* These results replicate Figure 8 for the full, including non-voting population.

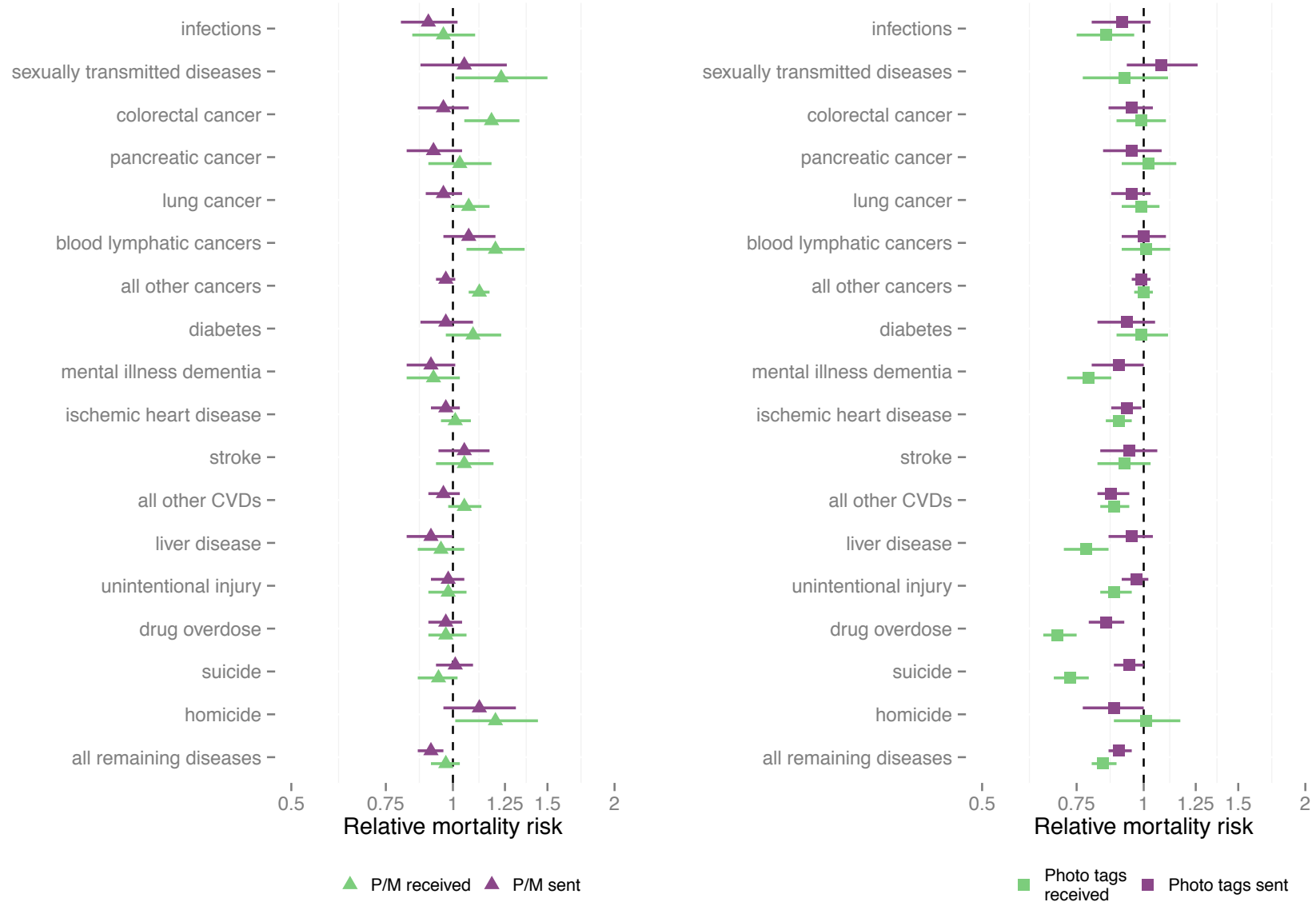

Fig. 14. *Facebook activities, by directionality and activity type (seventeen cause of death categories).* These results replicate Figure 9 for the full, including non-voting population.

L. Results for 'voter' population with controls added for 'full' population analyses

To more clearly compare the 'non-voter' population results to those in the main paper, we display here the 'voter' subsample with the additional non-voter population controls. We note that the friend-based controls were related to having more than just a few Facebook friends, and these correlations may be artifacts of the friend-based control method. We do not include them in the main paper results for this reason.

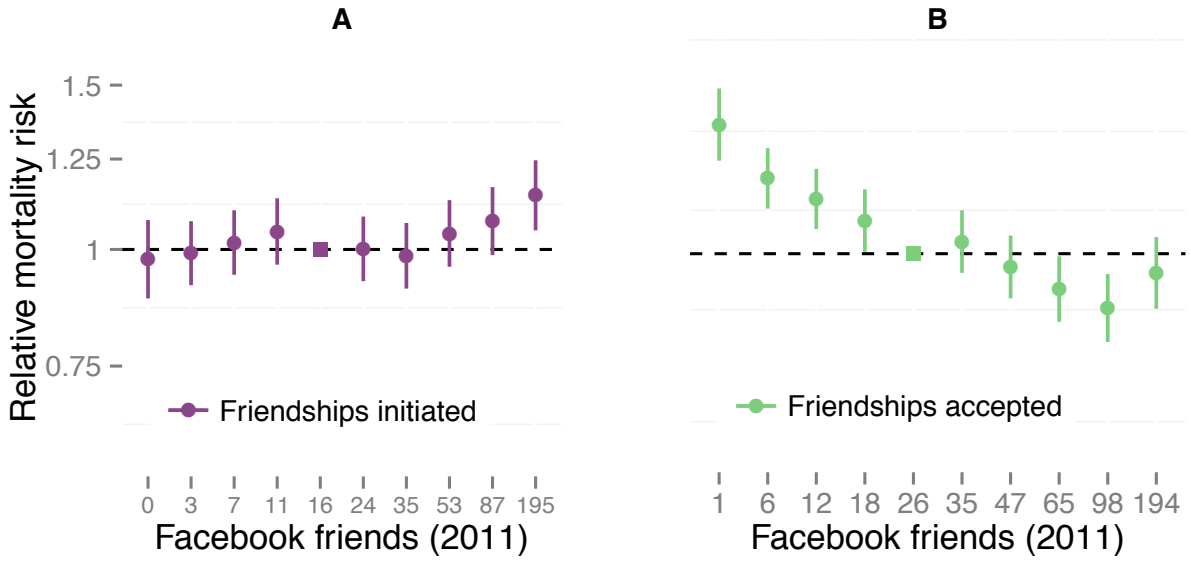

Fig. 15. *Facebook friends and relative mortality risk (all-cause mortality).* This figure shows all-cause mortality regressed on deciles of Facebook friend counts, by initiated and received Facebook friendships. These results replicate Figure 1 on the 'voting' population using the controls added for the full, non-voting population results.

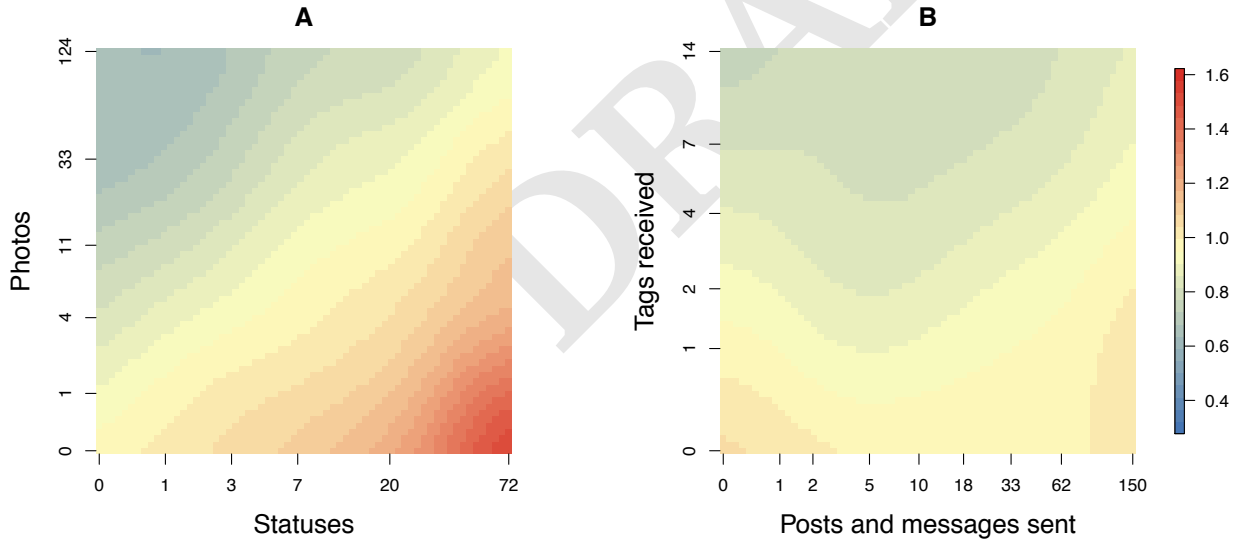

Fig. 16. *Facebook activities, smoothed mortality risk by quantile activity-activity interaction.* This figure shows all-cause mortality risk among Facebook users for combinations of activities. These results replicate Figure 2 on the 'voting' population using the controls added for the full, non-voting population results.

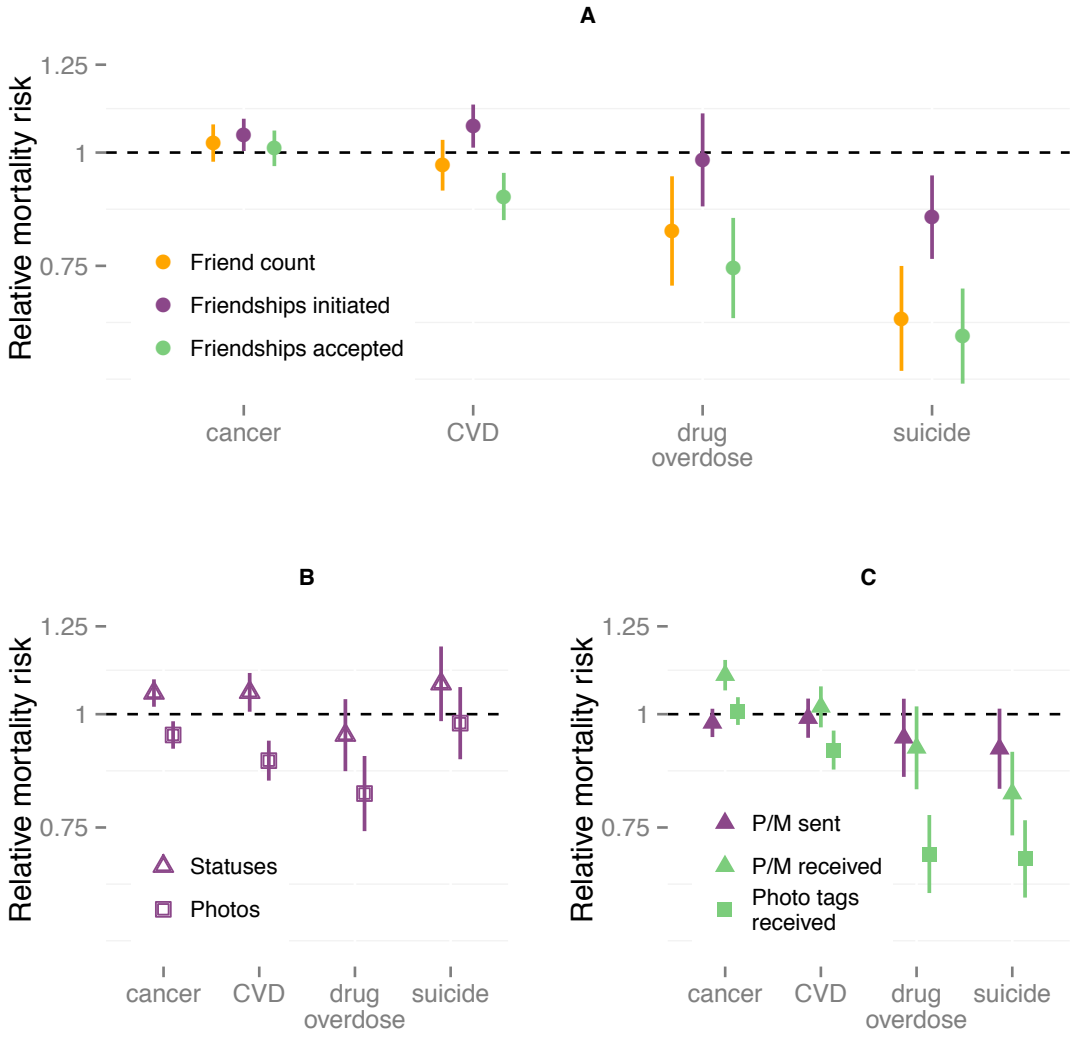

Fig. 17. *Facebook activities, by directionality and activity type.* This figure shows cause-specific morality risk among Facebook users by online activity levels (counts of Facebook tool use), separated by the directionality of the action (from or to the subject) and activity type (the type of Facebook tool used). These results replicate Figure 3 on the 'voting' population using the controls added for the full, non-voting population results.

M. Results for 'voter' population with race/ethnicity controls

As a robustness check, we repeated our main analyses with coarse controls for race/ethnicity. Race/ethnicity was inferred from last names, based on distributions in the US Census [8]. This method has previously been used to infer race/ethnicity of Twitter [9] and Facebook [10] users, in addition to common usage in health studies, e.g., cancer incidence [11] and health disparities [12]. We included dummy variables in our models for last names that were associated with 50% or more of US Census respondents reporting being "Hispanic", "Asian/Pacific Islander", or "Black". For last names not in the US Census data (approximately 12% of the voter record) or for which there was only partial race/ethnicity information, we imputed race/ethnicity using multivariate imputation by chained equations [13]. Our estimates for"Asian/Pacific Islander" and "Black" were lower than those seen in surveys of California registered voters [14]. According to this last name method, the California voter record as 64% white, 24% Hispanic, 6% Asian/Pacific Islander, and 1% black, while individuals on Facebook in Fall 2010 and in the California voter record were 70% white, 18% Hispanic, 6% Asian/Pacific Islander, and 1% black. These additional controls had small effects on our comparison of voters on Facebook vs. voters not on Facebook: After controlling for race/ethnicity, rates of suicide were slightly lower in the Facebook population than in the general population, though the effect was not robust to a Bonferonni correction, and race/ethnicity had no effect on our within-Facebook comparisons.

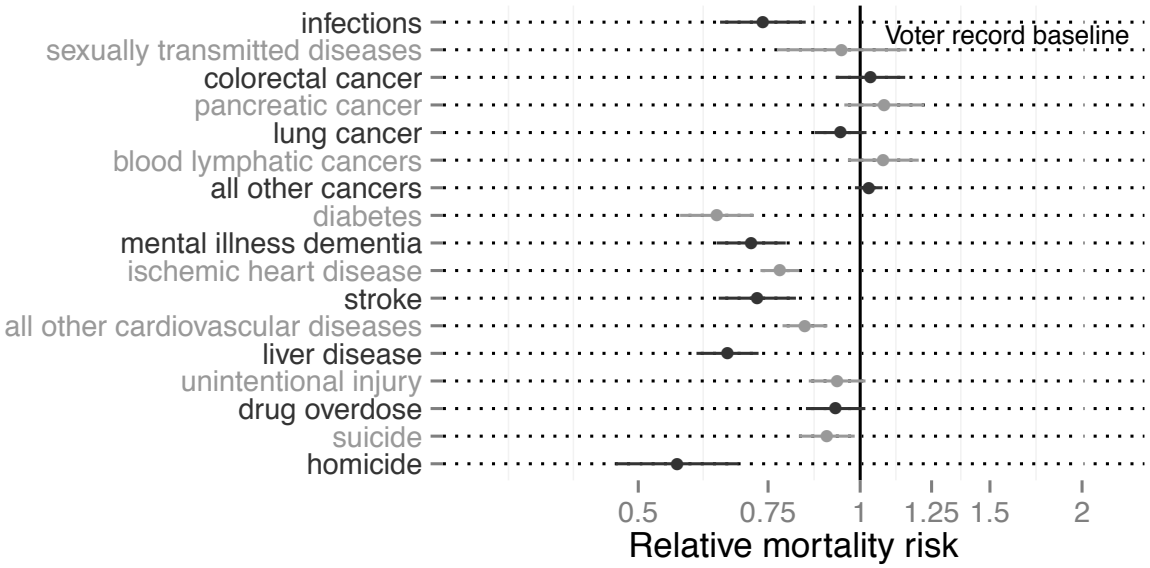

Fig. 18. *Facebook, California voter record mortality rate by cause comparison – with race/ethnicity control.* This figure shows relative cause-specific mortality risk among Facebook users whose first name, last name, and date of birth are listed in the California voter record compared to all California registered voters. The y-axis (on bottom) is the relative mortality risk estimated in a Cox proportional hazard model, and the x-axis is the cause of death. The all registered voter comparison group was exactly matched on age and gender so that the comparison groups are, by-design, perfectly balanced on these covariates.

Table 34. Infections risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.738
	(0.646, 0.844)
Male	1.689
	(1.512, 1.888)
Hispanic	1.382
	(1.215, 1.572)
Asian/Pacific Islander	0.425
	(0.300, 0.602)
Black Non-Hispanic	1.036
	(0.553, 1.940)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 35. Sexually transmitted diseases risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.944
	(0.771, 1.156)
Male	9.197
	(6.903, 12.253)
Hispanic	0.899
	(0.717, 1.128)
Asian/Pacific Islander	0.442
	(0.265, 0.739)
Black Non-Hispanic	1.568
	(0.710, 3.460)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 36. Colorectal cancer risk, voter record.

	Dependent variable:
	Deceased
Facebook	1.033
	(0.927, 1.151)
Male	1.278
	(1.159, 1.409)
Hispanic	0.894
	(0.784, 1.019)
Asian/Pacific Islander	0.735
	(0.580, 0.932)
Black Non-Hispanic	0.966
	(0.556, 1.677)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 37. Pancreatic cancer risk, voter record.

	Dependent variable:
	Deceased
Facebook	1.078
	(0.952, 1.221)
Male	1.482
	(1.323, 1.660)
Hispanic	0.952
	(0.817, 1.109)
Asian/Pacific Islander	0.703
	(0.530, 0.932)
Black Non-Hispanic	0.478
	(0.186, 1.228)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 38. Lung cancer risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.941
	(0.868, 1.020)
Male	1.140
	(1.061, 1.225)
Hispanic	0.541
	(0.480, 0.610)
Asian/Pacific Islander	0.684
	(0.573, 0.817)
Black Non-Hispanic	1.586
	(1.150, 2.188)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 39. Blood lymphatic cancers risk, voter record.

	Dependent variable:
	Deceased
Facebook	1.075
	(0.962, 1.201)
Male	1.561
	(1.412, 1.727)
Hispanic	1.133
	(1.000, 1.284)
Asian/Pacific Islander	0.768
	(0.605, 0.976)
Black Non-Hispanic	0.935
	(0.518, 1.687)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 40. All other cancers risk, voter record.

	Dependent variable:
	Deceased
Facebook	1.027
	(0.984, 1.072)
Male	0.923
	(0.887, 0.960)
Hispanic	0.926
	(0.879, 0.975)
Asian/Pacific Islander	0.775
	(0.706, 0.851)
Black Non-Hispanic	1.207
	(0.991, 1.469)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 41. Diabetes risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.639
	(0.570, 0.718)
Male	1.714
	(1.563, 1.879)
Hispanic	1.336
	(1.201, 1.487)
Asian/Pacific Islander	0.471
	(0.360, 0.617)
Black Non-Hispanic	1.616
	(1.068, 2.445)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 42. Mental illness dementia risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.711
	(0.639, 0.793)
Male	1.443
	(1.318, 1.579)
Hispanic	0.646
	(0.568, 0.736)
Asian/Pacific Islander	0.489
	(0.380, 0.629)
Black Non-Hispanic	0.854
	(0.514, 1.420)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 43. Ischemic heart disease risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.778
	(0.732, 0.827)
Male	3.039
	(2.871, 3.217)
Hispanic	0.771
	(0.717, 0.829)
Asian/Pacific Islander	0.566
	(0.495, 0.648)
Black Non-Hispanic	1.394
	(1.090, 1.783)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 44. Stroke risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.725
	(0.643, 0.818)
Male	1.277
	(1.156, 1.409)
Hispanic	1.143
	(1.010, 1.293)
Asian/Pacific Islander	1.081
	(0.881, 1.326)
Black Non-Hispanic	2.222
	(1.504, 3.283)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 45. all other CVDs risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.842
	(0.785, 0.903)
Male	2.002
	(1.880, 2.131)
Hispanic	0.731
	(0.672, 0.796)
Asian/Pacific Islander	0.425
	(0.354, 0.510)
Black Non-Hispanic	1.378
	(1.044, 1.819)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 46. Liver disease risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.661
	(0.600, 0.728)
Male	1.883
	(1.739, 2.038)
Hispanic	1.265
	(1.155, 1.384)
Asian/Pacific Islander	0.298
	(0.224, 0.397)
Black Non-Hispanic	0.896
	(0.565, 1.422)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 47. Unintentional injury risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.931
	(0.852, 1.017)
Male	3.217
	(2.941, 3.518)
Hispanic	0.845
	(0.767, 0.931)
Asian/Pacific Islander	0.479
	(0.385, 0.595)
Black Non-Hispanic	0.976
	(0.650, 1.465)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 48. Drug overdose risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.926
	(0.844, 1.016)
Male	1.657
	(1.523, 1.803)
Hispanic	0.542
	(0.483, 0.609)
Asian/Pacific Islander	0.224
	(0.162, 0.309)
Black Non-Hispanic	1.091
	(0.747, 1.594)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 49. Suicide risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.901
	(0.826, 0.983)
Male	3.083
	(2.824, 3.365)
Hispanic	0.465
	(0.414, 0.522)
Asian/Pacific Islander	0.565
	(0.467, 0.683)
Black Non-Hispanic	0.935
	(0.633, 1.380)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

Table 50. Homicide risk, voter record.

	Dependent variable:
	Deceased
Facebook	0.564
	(0.466, 0.684)
Male	3.411
	(2.874, 4.047)
Hispanic	1.193
	(1.016, 1.401)
Asian/Pacific Islander	0.580
	(0.393, 0.855)
Black Non-Hispanic	2.525
	(1.558, 4.092)
Observations	11,834,575
Note:	95% confidence intervals in parentheses.

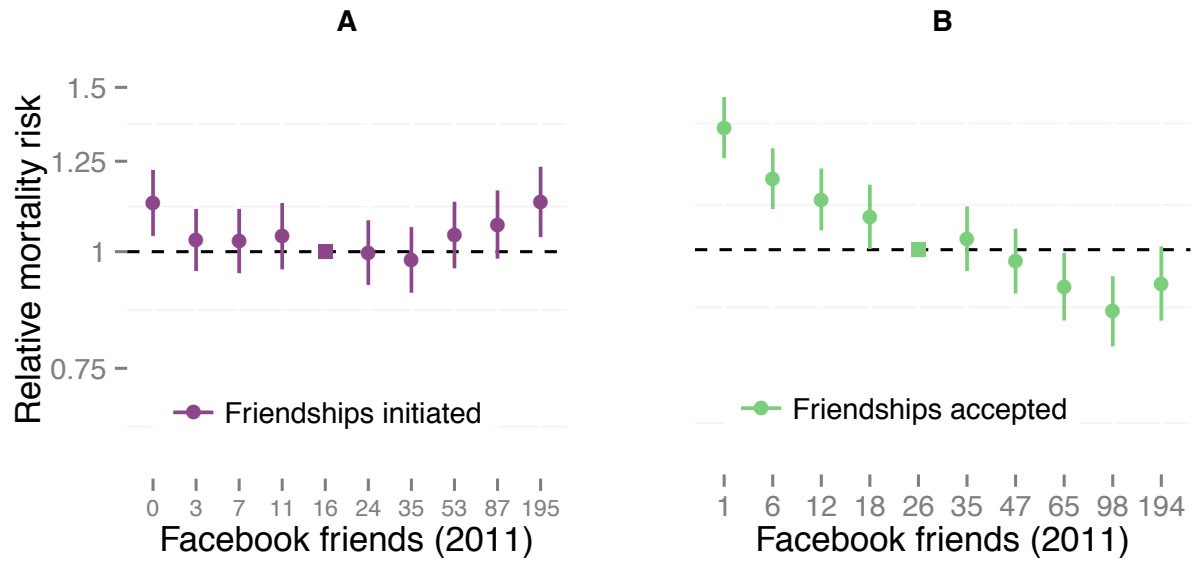

Fig. 19. *Facebook friends and relative mortality risk (all-cause mortality) – with race/ethnicity control.* This figure shows all-cause mortality regressed on deciles of Facebook friend counts, by initiated and received Facebook friendships. These results replicate Figure 1 on the 'voting' population using the controls added for the full, non-voting population results.

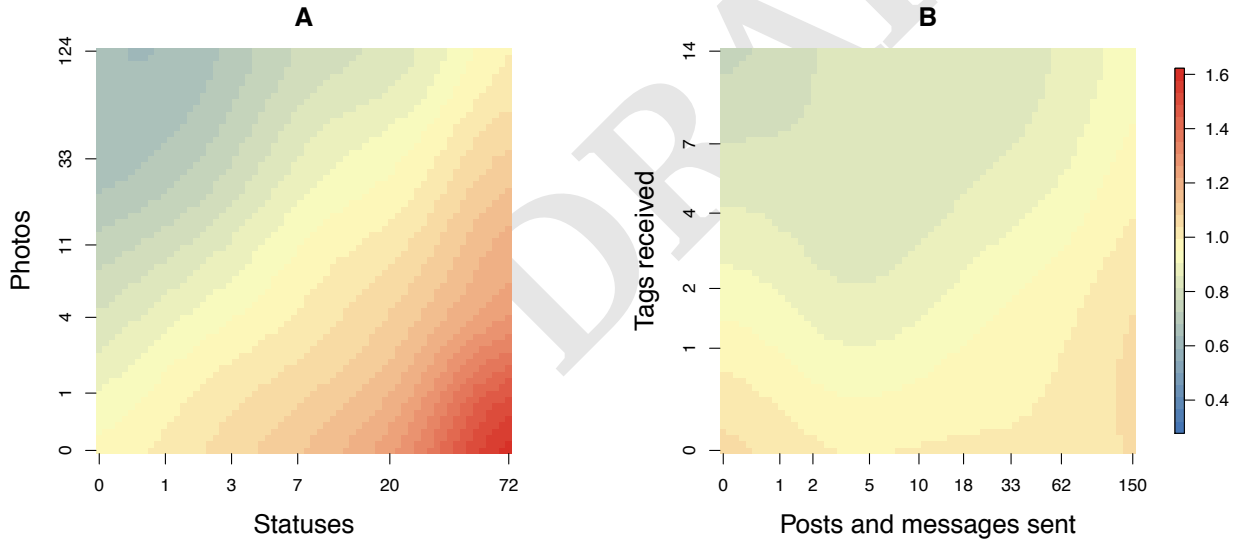

Fig. 20. *Facebook activities, smoothed mortality risk by quantile activity-activity interaction – with race/ethnicity control.* This figure shows all-cause mortality risk among Facebook users for combinations of activities. These results replicate Figure 2 on the 'voting' population using the controls added for the full, non-voting population results.

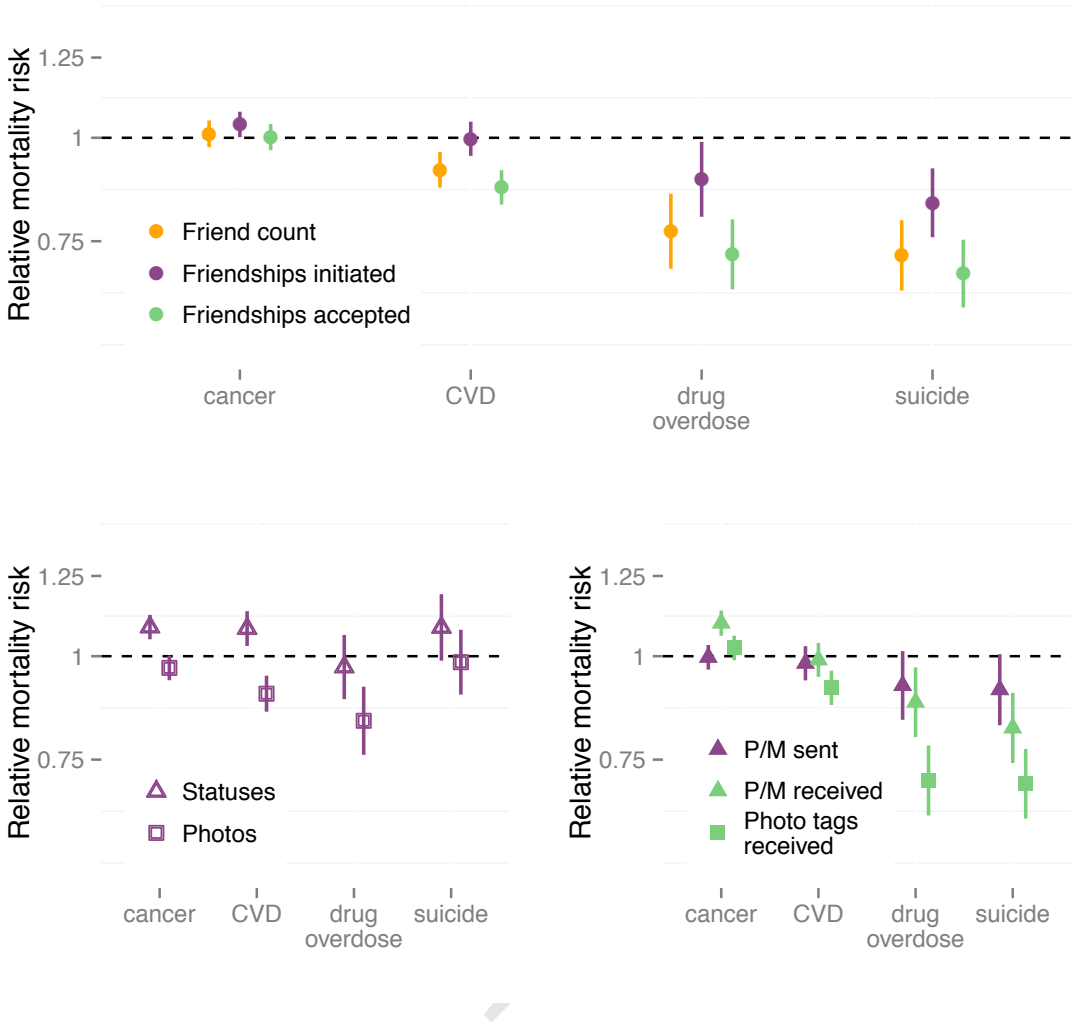

Fig. 21. *Facebook activities, by directionality and activity type – with race/ethnicity control.* This figure shows cause-specific morality risk among Facebook users by online activity levels (counts of Facebook tool use), separated by the directionality of the action (from or to the subject) and activity type (the type of Facebook tool used). These results replicate Figure 3 on the 'voting' population using the controls added for the full, non-voting population results.

1. Duggan M, Ellison NB, Lampe C, Lenhart A, Madden M (2015) Social Media Update 2014, (Pew Research Center, Washington, DC), Technical report.
2. Bond R, Fariss C, Fowler JH, Jones J, Marlow C (2012) A 61-million-person experiment in social influence and political mobilization. *Nature* 489:295–298.
3. Thorp AA, Owen N, Neuhaus M, Dunstan DW (2011) Sedentary Behaviors and Subsequent Health Outcomes in Adults. *American Journal of Preventative Medicine* 41(2):207–215.
4. Hampton KN, Rainie L, Lu W, Shin I, Purcell K (2015) Social Media and the Cost of Caring, (Pew Research Center, Washington, DC), Technical report.
5. Christakis NA, Allison PD (2006) Mortality after the hospitalization of a spouse. *New England Journal of Medicine* 354(7):719–730.
6. Smith A (2011) 35% of American adults own a smartphone, (Pew Research Center, Washington, DC), Technical report.
7. Jolliffe I (2002) *Principal component analysis*. (Wiley Online Library).
8. US Census Bureau (2000) Genealogy data: Frequently occurring surnames from census 2000. . *http://www.census.gov/topics/population/genealogy/data/2000_surnames.html*.
9. Mislove A, Lehmann S, Ahn YY, Onnela JP (2011) Understanding the Demographics of Twitter Users. *ICWSM*.
10. Chang J, Rosenn I, Backstrom L, Marlow C (2010) ePluribus: Ethnicity on Social Networks. *ICWSM*.
11. Yost K, Perkins C, Cohen R, Morris C, Wright W (2001) Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. *Cancer Causes & Control* 12(8):703–711.
12. Elliott MN et al. (2009) Using the Census Bureau's surname list to improve estimates of race/ethnicity and associated disparities. *Health Services and Outcomes Research Methodology* 9(2):69–83.
13. Buuren S, Groothuis-Oudshoorn K (2011) mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software* 45(3).
14. Baldassare M, Bonner D, Kordus D, Lopes L (2015) Voter Participation in California, (Public Policy Institute of California), Technical report.