nature portfolio

Peer Review File

Positive association between Internet use and mental health among adults aged ≥50 years in 23 countries

Corresponding Author: Professor Qingpeng Zhang

This manuscript has been previously reviewed at another journal that is not operating a transparent peer review scheme. The manuscript was considered suitable for publication without further review at Nature Human Behaviour.

This file contains all editorial decision letters in order by version, followed by all author rebuttals in order by version.

Version 0:

Decision Letter:

26th March 2024

Dear Professor Zhang,

Thank you once again for your manuscript, entitled "The beneficial impact of Internet use on the mental well-being among adults aged ≥50 years in 23 countries", and for your patience during the peer review process.

Your Article has now been evaluated by 3 referees. You will see from their comments copied below that, although they find your work of considerable potential interest, they have raised quite substantial technical and conceptual concerns. In light of these comments, we cannot accept the manuscript for publication, but would be interested in considering a revised version if you are willing and able to fully address reviewer and editorial concerns.

We hope you will find the referees' comments useful as you decide how to proceed. If you wish to submit a substantially revised manuscript, please bear in mind that we will be reluctant to approach the referees again in the absence of major revisions. All 3 reviewers raise technical concerns, which we expect to see addressed in full in your revised manuscript, with further analyses added as requested. Please also rewrite your manuscript as needed to address our reviewers' conceptual concerns about the theoretical motivation and interpretation of your work.

We are committed to providing a fair and constructive peer-review process. Do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

Finally, your revised manuscript must comply fully with our editorial policies and formatting requirements. Failure to do so will result in your manuscript being returned to you, which will delay its consideration. To assist you in this process, I have attached a checklist that lists all of our requirements. If you have any questions about any of our policies or formatting, please don't hesitate to contact me.

If you wish to submit a suitably revised manuscript, we would hope to receive it within 3 months. I would be grateful if you could contact us as soon as possible if you foresee difficulties with meeting this target resubmission date.

With your revision, please:

• Include a “Response to the editors and reviewers” document detailing, point-by-point, how you addressed each editor and referee comment. If no action was taken to address a point, you must provide a compelling argument. When formatting this document, please respond to each reviewer comment individually, including the full text of the reviewer comment verbatim followed by your response to the individual point. This response will be used by the editors to evaluate your revision and sent back to the reviewers along with the revised manuscript.

• Highlight all changes made to your manuscript or provide us with a version that tracks changes.

Please use the link below to submit your revised manuscript and related files:

Link Redacted

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

Thank you for the opportunity to review your work. Please do not hesitate to contact me if you have any questions or would like to discuss the required revisions further.

Sincerely,

[redacted]
Senior Editor
Nature Human Behaviour

Reviewer expertise:

Reviewer #1: internet use in older adults; mental health/wellbeing in older adults; public health

Reviewer #2: internet use in older adults; mental health/wellbeing in older adults; public health

Reviewer #3: mental health/wellbeing in older adults; public health

REVIEWER COMMENTS:

Reviewer #1:
Remarks to the Author:
This study investigates the relationship between internet use and mental well-being outcomes (depressive symptoms, life satisfaction, and self-reported health) across aging cohorts from 23 countries. The novelties include:

1. Assessing the frequency of internet use, in addition to binary measures of use/non-use.
2. Evaluating potential effect heterogeneity by sociodemograpsohic characteristics and genetic risk factors.

The data harmonization and analysis across these numerous international cohorts represents a substantial undertaking that should be commended. However, there are some significant concerns regarding the presentation and methodology that need to be addressed:

Major Comments

1. The authors need to provide more theoretical backgrounds to motivate their study questions. In the Introduction, the authors claim it is important to consider sociocultural contexts when considering the internet-wellbeing associations, which I appreciate. However, they conducted a meta-analysis across countries and provided a lot of discussion based on the meta-analyzed relationship. Why was a meta-analysis necessary? In particular, given that the observed associations were in the opposite direction in some countries, pooling estimates across countries would not provide meaningful insights. Similarly, in the first paragraph of the Introduction, they wrote about the importance of addressing mental health problems, which is a nice justification for examining depressive symptoms as an outcome. Why did they look at life satisfaction and self-rated health as outcomes too? They used the term "mental well-being" as an umbrella term to capture all three outcomes. However, mental illness and subjective well-being are conceptually different. Self-rated health does not necessarily reflect one's mental health and is also affected by somatic symptoms.

2. A related issue is the lack of motivating statements regarding policy implications of estimating heterogeneous effects, particularly by genetic risks. It is certainly true that effects may vary by sociodemographic and genetic characteristics. Why is it so important to look at such heterogeneity? Does it improve the way we promote mental health in a meaningful manner? In clinical settings where "precision-medicine" interventions are feasible, it may be useful to have a better understanding of effect heterogeneity. The authors briefly mentioned the importance of tailored digital inclusion. What is it? Is it feasible to realistically implement in policy settings? I think the manuscript overall lacks clarity on how the findings can be useful from policy/clinical perspectives.

3. For the main analysis, many aspects of their analytic approach remained unclear after reading what's written in the manuscript. The authors wrote that they used multiple waves of data and modeled it with mixed-effects regression with years since baseline as the timescale. This suggests that they specified a functional form (presumably linear) for the time trend in the outcome and added the internet use variable as a separate term. This model effectively assumes that the association between internet use and outcomes does not vary over time. This is an assumption that is hard to believe and in fact contradicts their own empirical findings about the effect heterogeneity by age. Providing an equation for the regression model would be appreciated. It was also unclear why the authors provided the joint effect analysis only for genes. Unless there are ways to intervene and modify genes, the interaction analysis should suffice for genes as well. This way, the

authors can combine the results for all other sociodemographic characteristics with those for genes.

4. The authors analyzed those who responded to all waves. This means that the sample selection occurred after the baseline treatment (i.e., the internet use), leading to selection bias (also known as collider stratification bias). They further removed individuals with missing data, further increasing the risk of selection bias. For the former point, the authors are encouraged to compare baseline characteristics between the whole baseline participants (e.g., wave 9 participants in HRS) and the analytic samples. If the attrition is associated with the treatment status and other covariates that are potentially associated with the outcomes, then selection bias is going to be a concern. For the second point, the authors showed results based on multiple imputation. I think it makes more sense to show these findings as the main results rather than showing them as sensitivity analyses.

5. Their analysis of cumulative internet use is essentially flawed because it ignores the presence of time-dependent confounders. It is well known that the estimation of a sustained treatment (e.g., cumulative internet use) requires more sophisticated approaches for confounding adjustment (e.g., g-computation).

6. Discussions of the findings should be more balanced. The discussions are generally based on the premise that the internet is good for mental well-being, and this premise is reflected in the emphasis on interpreting the meta-analyzed results and the conclusion highlighting the positive effects. The results seem to be more nuanced than that, with some results even showing adverse effects. Also, the authors' discussion of policy implications was mostly about China. I encourage the authors to expand that to more global contexts.

7. Some statements are not empirically supported. First, the authors emphasized their findings about dose-response relationships. However, looking at the figures, I do not see strong evidence of dose-response relationships except for the effect on depressive symptoms and self-reported health in the USA. Second, the conclusion emphasizes the importance of enhancing digital and health literacies. This is a stretch to say and not empirically supported by their findings. The authors also mentioned the Sustainable Development Goals (SDGs) in the Conclusion without providing context in the Introduction, which I thought was strange.

Minor comments

1. I appreciate their efforts to conduct many analyses to get nuanced results and report them using a unique data visualization approach. However, the figures with tons of information are overwhelming and hard to digest for most readers, and these results are mostly not discussed anyway. I suggest the authors present more condensed results in the main text and show these granular findings as Supplement.

2. The analysis of country-level indicators (e.g., Gini coefficient) seems unnecessary. These are purely exploratory analyses with strong modeling assumptions. The high R-squared values do not suggest the importance of the indicator itself. Also, it was unclear why they chose to examine correlations for clearly non-linear and non-monotonic relationships.

3. The authors need to theoretically justify their choice of cut-offs for effect modifiers (e.g., below vs above 65 for age).

4. In Results, the authors should avoid interpretative and subjective language. For example, the authors used the term “effect” to refer to statistical associations. I believe the authors know the distinction between these two concepts but being precise on such a language will improve the manuscript. Similarly, the authors used the word “significantly” to refer to the effect sizes rather than statistical significance. This is a very subjective and vague word, and should be avoided.

Reviewer #2:
Remarks to the Author:
This is a nice written manuscript investigated the association of internet use and depressive symptoms in people aged 50 and over from six aging cohorts. The association was consistent across the varied cohorts, and some sensitivity analyses also support the conclusion. Here are some comments for consideration.
1. Did the authors analyze the data with reversing the x (depressive symptoms) and y (internet use)? Is this association could be bi-directional?
2. Did the authors try to investigate the association of internet use and incident of depressive symtoms (none at baseline but occur in the follow-ups)?
3. I have come across several highly cited paper which have already reported the very topic, what's the extra contribution, in addition to multi cohorts, for this one?
Cotten S R, Ford G, Ford S, et al. Internet use and depression among older adults[J]. Computers in human behavior, 2012, 28(2): 496-499.
Cotten S R, Ford G, Ford S, et al. Internet use and depression among retired older adults in the United States: A longitudinal analysis[J]. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 2014, 69(5): 763-771.
Liao S, Zhou Y, Liu Y, et al. Variety, frequency, and type of Internet use and its association with risk of depression in middle- and older-aged Chinese: a cross-sectional study[J]. Journal of Affective Disorders, 2020, 273: 280-290.
4. The authors have discussed the mechanisms between the internet use and depression, family connection, which is good, do we consider to control it in the regression and do subgroup analyses? In view of your explanation, the associaiton could be more marked in those who contact family members or friends less frequently.

Reviewer #3:
Remarks to the Author:
The paper provides a good narrative on the association between Internet use and age-related mental health challenges in older adults. The data from six aging cohorts across 23 high- and middle-income countries is well-defined.
As the authors have described, the protective role of Internet use varies in different subpopulations within a country, an important factor to consider when reviewing the social determinants of health. However, we need to consider the question of whether a higher frequency of Internet use is related to improved mental health. These findings could provide significant implications for public health policies and practices in promoting mental well-being in later life through Internet usage. This paper is well-written and contains useful information. I have a few minor points the authors may wish to consider.

1. The authors should consider reporting why they chose these datasets only.
2. Include internet usage as social support for older adults in measurable terms
3. Explain the negative association (i.e., more beneficial) between Internet use and depressive symptoms in Fig.3a.
4. Sample selection procedures for each cohort need further clarification in Fig.1.
5. In the introduction section, one would expect a coherent exposition on well-being defined as high and low, moderate, or medium.
6. Further explain the meta-regression analyses on the positive impact of Internet use on self-reported health and how it increased with the GDP per capita and world happiness index at the country level.

Version 1:

Decision Letter:

Our ref: NATHUMBEHAV-24010344A

29th July 2024

Dear Dr. Zhang,

Thank you for submitting your revised manuscript "The beneficial impact of Internet use on the mental well-being among adults aged ≥50 years in 23 countries" (NATHUMBEHAV-24010344A). It has now been seen by the original referees and their comments are below. As you can see, the reviewers find that the paper has improved in revision. We will therefore be happy in principle to publish it in Nature Human Behaviour, pending minor revisions to satisfy the referees' final requests and to comply with our editorial and formatting guidelines. Please note that this will include requests that you remove the emphasis on the longitudinal nature of your analyses given the lack of temporal ordering, use the correct approach for pooling estimates after multiple imputation (CIs should be pooled using Rubin's rule), and remove causal language from the manuscript.

We are now performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements within two weeks. Please do not upload the final materials and make any revisions until you receive this additional information from us.

Please do not hesitate to contact me if you have any questions.

Sincerely,

[redacted]
Senior Editor
Nature Human Behaviour

Reviewer #1 (Remarks to the Author):

The authors have done many additional analyses and revisions to address my previous comments, which I appreciate. Upon reading the revised version, some concerns remain.

1. The addition of the equation for the main analysis is helpful. It seems that the authors are essentially looking at cross-sectional associations between internet use and the outcomes. In their "longitudinal" model, internet use was assessed at the same time point as the outcome assessment. The model allowed this association to vary over time (via the interaction term with time). The time-updated cross-sectional associations were then pooled as AME. Hence, much of the observed associations are likely due to reverse causation due to the lack of temporal ordering between the internet use and the

outcomes. The authors did a sensitivity analysis excluding those with baseline depression, but this will address reverse causation for the "baseline" internet use only. Subsequent cross-sectional associations remain subject to reverse causation. Nonetheless, the paper overall emphasized the longitudinal nature of the analysis.

2. I still believe the country-level factors analysis is uninformative and unnecessary. It is purely descriptive, yet the authors interpreted the meta-regression results as if the between-country variation can be attributed to each specific country-level indicator. They used language such as "attributed" or "explain". Identifying the source of the between-country heterogeneity would be a whole another paper, and I suggest the authors keep the paper more focused on identifying the heterogeneity itself. The paper is already full of information, and adding extra information that is very exploratory obscures the key contribution of the study (and it is not reader-friendly).

3. The authors are using an incorrect approach for pooling estimates after multiple imputation. Specifically, confidence intervals need to be pooled via Rubin's rule and are not simple mean of lower/upper bounds from each imputed dataset.

Reviewer #2 (Remarks to the Author):

The authors have improved the manuscript substantially and I have no further comments.

Dear Editor and Reviewers,

Thank you so much for reviewing our paper! We really appreciate the helpful comments and constructive suggestions! We have substantially revised the manuscript according to the review comments. All revised portions are marked in red in the revised manuscript.

Specific comments and our responses are detailed below.

REVIEWER 1'S COMMENTS

Remarks to the Author:

COMMENT 1-1. *The authors need to provide more theoretical backgrounds to motivate their study questions. In the Introduction, the authors claim it is important to consider sociocultural contexts when considering the internet-wellbeing associations, which I appreciate. However, they conducted a meta-analysis across countries and provided a lot of discussion based on the meta-analyzed relationship. Why was a meta-analysis necessary? In particular, given that the observed associations were in the opposite direction in some countries, pooling estimates across countries would not provide meaningful insights.*

RESPONSE 1-1. Thank you very much for the thoughtful comments and suggestions. We acknowledge the importance of providing more theoretical backgrounds for the motivation behind our study. Following the reviewer's suggestion, we added the justifications for the use of meta-analysis in the **Introduction** Section. Meta-analysis is essential in our study because it allows us to gain a whole picture of whether Internet use is effective in improving mental health for further public health decision-making using evidence from multiple countries (Riley et al., 2010). Contradictory results are common in the meta-analysis especially with continuous outcomes (Alba et al., 2016). Such heterogeneity, based on data or study design, can be addressed by the random-effects meta-analysis, and explained by the subgroup analysis or meta-regression (Alba et al., 2016; Borenstein et al., 2021). Notably, our subgroup meta-analyses (**Supplementary Table 9**) showed significant differences in the average marginal effect between studies with different measures of Internet use. Because the measure of Internet use is a binary indicator (*use* or *not use*) in most countries, we repeated the analysis after excluding participants from England and Mexico, where measures are not binary. The results are consistent. Moreover, we conducted meta-regressions (**Supplementary Fig.6**) with seven country-level factors to assess how much the heterogeneity can be explained. These analyses indicate that our findings are robust and provide several insights on cross-national variations. We recognize the limitations of pooling estimates across countries, especially when associations are in the opposite direction in some countries. Therefore, we provided more discussions on the non-protective association between Internet use and mental health.

The revised parts are copied below:

In the main text:

Introduction Section, **Fourth** Paragraph: “IPD meta-analysis allows for a synthesis of evidence from multiple cohorts to establish an overall summary of whether Internet use is effective in improving mental health for further public health decision-making[36].”

Discussion Section, **Ninth and Tenth** Paragraphs: “Notably, we observed that the protective association between Internet use and mental health became non-protective in England (only for life satisfaction), and insignificant in half of other countries after adjusting for demographic characteristics, socioeconomic positions, health behaviors, and physical health status (Supplementary Fig.21–23 and Supplementary Tables 6–8). As we discussed above, the different roles of Internet use may reflect the disparities across countries, subpopulations, and even individuals. These findings highlight the need to examine the varying effect of Internet use across countries in different subpopulations within a country.

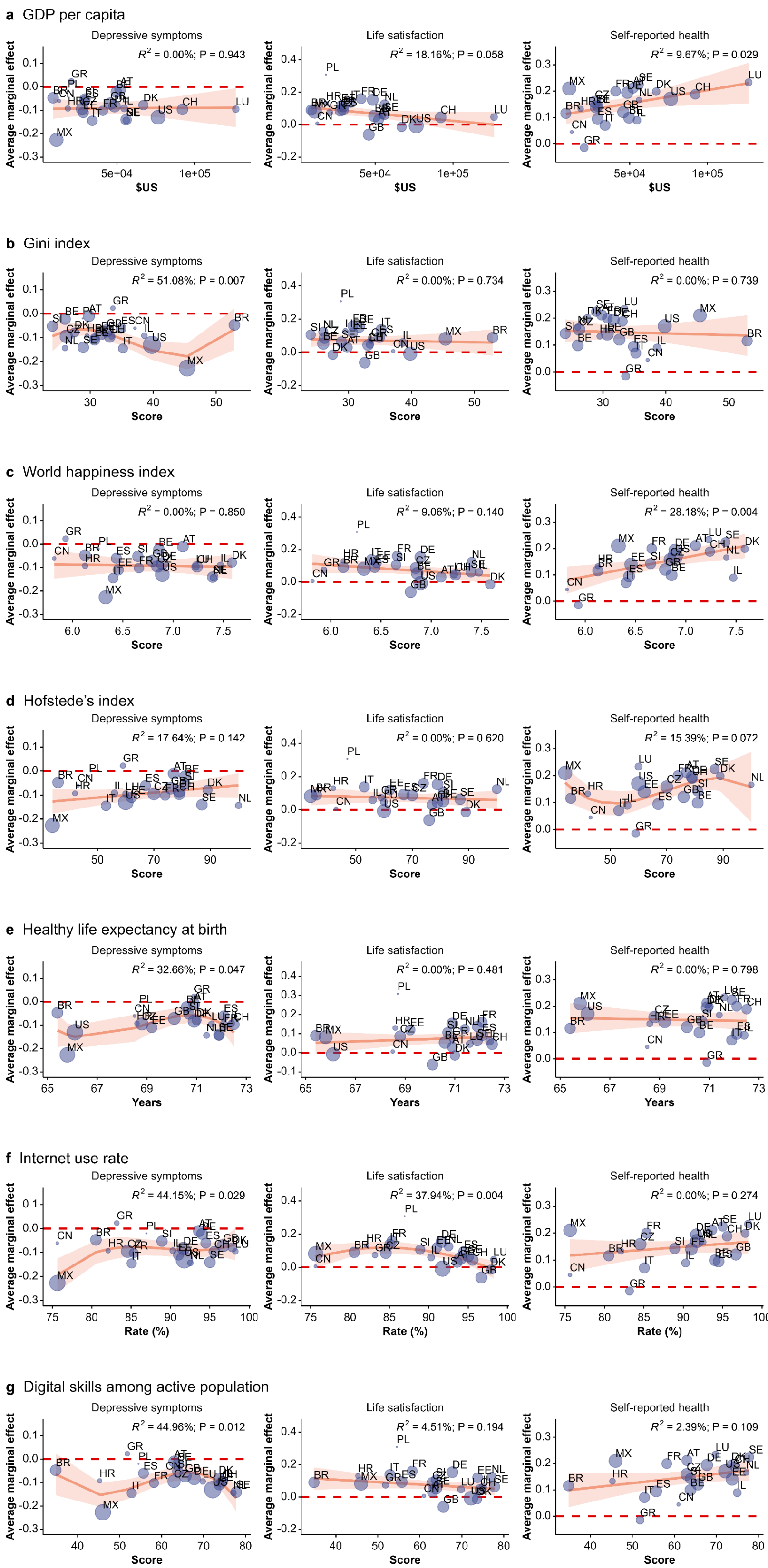
Nonetheless, Internet use may be a double-edged sword for mental health among older adults. For example, excessive online duration has found to be associated with poor mental health, possibly by interfering with other essential offline activities, including sleep, physical activity, and real-life social activity[22,23]. Additional Internet use-related risk factors for mental health include cyberbullying, online fraud, and misinformation[59–61]. However, these risk factors were not captured by our datasets. Future research should consider more details of Internet use, such as daily screen time, purposes of use, types of online activities, and the interactions with other users, to inform effective Internet use-related intervention strategies for mental health promotion[62].”

In the Supplementary Materials:

Supplementary Table 9: Subgroup analyses by Internet use measures in the meta-analysis.

Internet use measures	AME (95% CI)	I^2	τ^2	H^2	$P_{subgroup}$
Depressive symptoms					<0.001
Binary	-0.08 [-0.10, -0.06]	44.55	0.001	1.803	
Frequency-based	-0.07 [-0.12, -0.02]	0.00	0.000	1.00	
Household-level	-0.23 [-0.26, -0.20]	0.00	0.000	1.00	
Life satisfaction					0.029
Binary	0.08 [0.05, 0.10]	63.31	0.002	2.726	
Frequency-based	-0.06 [-0.12, -0.01]	0.00	0.000	1.00	
Household-level	0.08 [0.05, 0.11]	0.00	0.000	1.00	
Self-reported health					0.388
Binary	0.15 [0.12, 0.17]	70.93	0.002	3.440	
Frequency-based	0.12 [0.07, 0.17]	0.00	0.000	1.00	
Household-level	0.21 [0.18, 0.24]	0.00	0.000	1.00	

Note: Internet use measures were binary in HRS, SHARE, CHARLS, and ELSI, household-level in MHAS, and frequency-based in ELSA. AME = average marginal effect, CI = confidence interval, HRS = Health and Retirement Study, ELSA = English Longitudinal Study on Ageing, ELSI = Brazilian Longitudinal Study of Aging, SHARE = Survey of Health, Ageing and Retirement in Europe, CHARLS = China Health and Retirement Longitudinal Study, MHAS = Mexican Health and Aging Study.



Supplementary Fig.6: Bubble plot with fitted meta regression line of the average marginal effect for three mental health outcomes and country-level factors, including GDP per capita (a), Gini index (b), world happiness index (c), Hofstede's index (d), healthy life expectancy at birth (e), Internet use rate (f), and digital skills among active population (g). Data are presented as the average marginal effects (bubbles), 95% confidence intervals (shaded area), and the linear prediction (orange line). Red dashed line at zero represents the reference line. The R^2 quantifies the proportion of heterogeneity that can be accounted for by each country-level factor. The P value indicates whether the average marginal effect was associated with each country-level factor. AT = Austria, BE = Belgium, BR = Brazil, CN = China, HR = Croatia, CZ = Czech Republic, DK = Denmark, GB = England, EE = Estonia, FR = France, DE = Germany, GR = Greece, IL = Israel, IT = Italy, LU = Luxembourg, MX = Mexico, NL = Netherlands, PL = Poland, SI = Slovenia, ES = Spain, SE = Sweden, CH = Switzerland, and US = United States.

References:

Alba, A. C., Alexander, P. E., Chang, J., MacIsaac, J., DeFry, S., & Guyatt, G. H. (2016). High statistical heterogeneity is more frequent in meta-analysis of continuous than binary outcomes. *Journal of Clinical Epidemiology*, *70*, 129–135. https://doi.org/10.1016/j.jclinepi.2015.09.005

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to Meta-Analysis*. John Wiley & Sons.

Riley, R. D., Lambert, P. C., & Abo-Zaid, G. (2010). Meta-analysis of individual participant data: Rationale, conduct, and reporting. *BMJ*, *340*, c221. https://doi.org/10.1136/bmj.c221

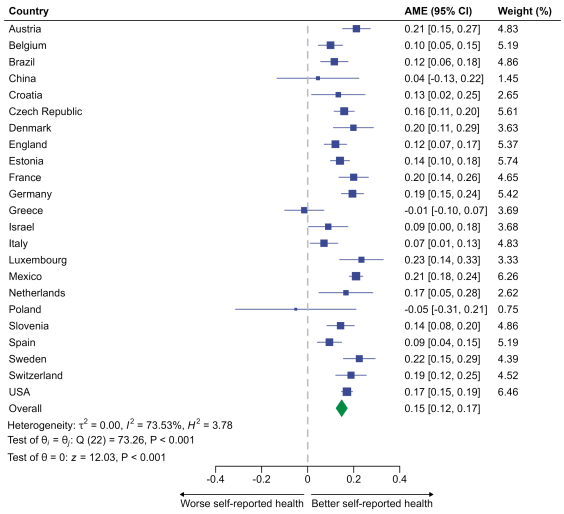
COMMENT 1-2. *Similarly, in the first paragraph of the Introduction, they wrote about the importance of addressing mental health problems, which is a nice justification for examining depressive symptoms as an outcome. Why did they look at life satisfaction and self-rated health as outcomes too? They used the term "mental well-being" as an umbrella term to capture all three outcomes. However, mental illness and subjective well-being are conceptually different. Self-rated health does not necessarily reflect one's mental health and is also affected by somatic symptoms.*

RESPONSE 1-2. Thank you so much for the comments. To justify the use of depressive symptoms and life satisfaction as outcomes, we first introduced the definition of mental health proposed by the World Health Organization (WHO) and the staging model for mental health developed by The Lancet Commission on Global Mental Health and Sustainable Development in the **Introduction** Section (World Health Organization, 2022; Patel et al., 2018). According to this staging model, mental health is a complex continuum covering from well-being (e.g., life satisfaction) to conditions (e.g., depressive symptoms). Therefore, it is essential to consider both depressive symptoms and life satisfaction when investigating the association between Internet use and mental health. According to the staging model, "mental health" refers to the full spectrum of mental health states, "mental illness" indicates pathological disease states, and "mental well-being" acts as the positive component of mental health. So, we replaced the term "mental well-being" with "mental health" in the revised manuscript. Finally, we agree with the reviewer that self-reported health is an indicator that reflects a person's integrated assessment of both physical and mental health status, rather than a traditional mental health indicator (Miilunpalo et al., 1997). Following the reviewer's suggestion, we moved the results of self-reported health from the manuscript to supplementary materials. Nonetheless, since self-reported health and mental health are highly correlated and the results regarding self-reported health are largely consistent with depressive symptoms and life satisfaction (Peleg & Nudelman, 2021), we retained the corresponding results and discussions. We copied the modifications in the revised manuscript below:

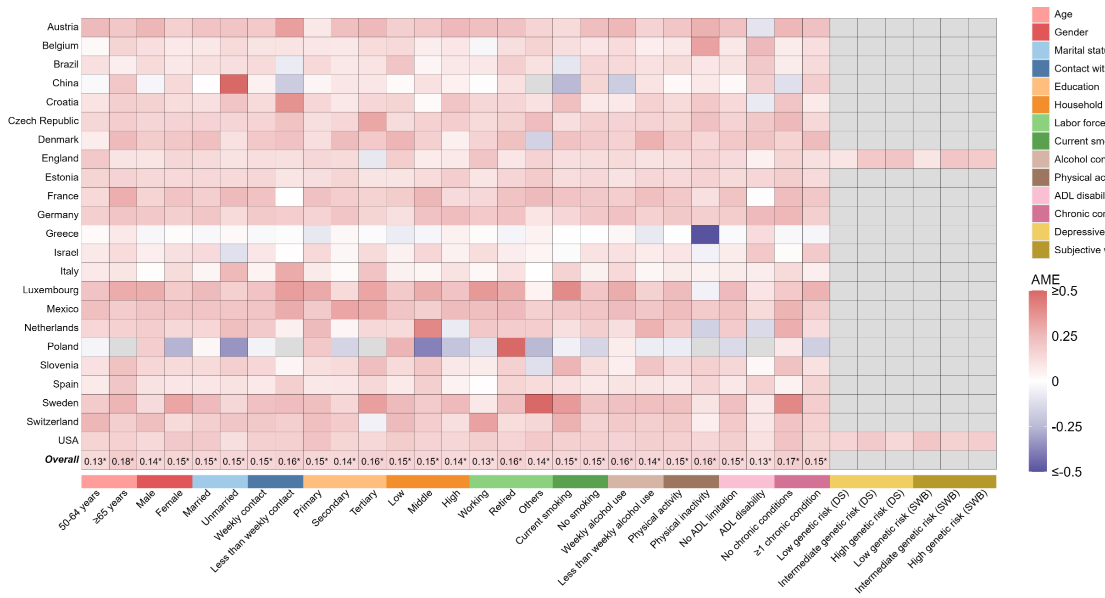
In the main text:

Introduction Section, **First** Paragraph: "Mental health, as defined by the World Health Organization (WHO), is a state of well-being that enables individuals to manage the stresses of life, realize their abilities, learn and work well, and contribute to their community[2]. This definition extends beyond the absence of mental disorders. According to the staging model for mental health developed by The Lancet Commission on Global Mental Health and Sustainable Development, mental health encompasses a broad spectrum, ranging from well-being (e.g., life satisfaction) to various conditions (e.g., depressive symptoms)[2,3]. These concepts advocate for mental health promotion, emphasizing the need to not only prevent mental disorders but also enhance mental well-being."

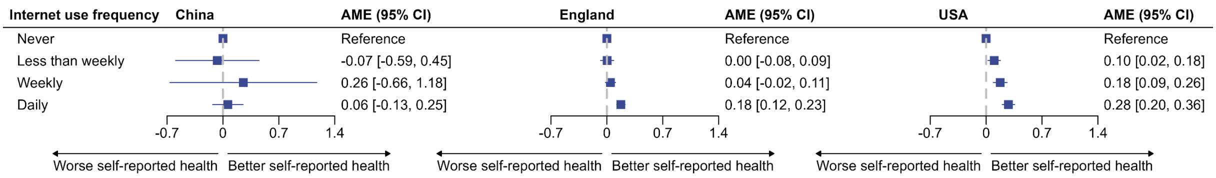
In the Supplementary Materials:



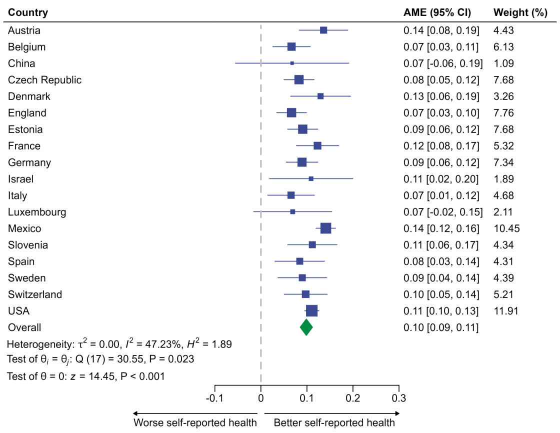
Supplementary Fig.1: Associations of Internet use with self-reported health. Data are presented as the AME coefficients with 95% CIs. The AME for each country (blue square) was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall AME (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.



Supplementary Fig.8: Associations of Internet use with self-reported health by age, gender, marital status, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and genetic risks assessed by PGS for depressive symptoms and subjective well-being. The AME for each country was separately estimated by the linear mixed model adjusted for all covariates (not including genetic risks) except the stratification variable. The overall AME was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. The color of cells indicates the effect sizes (AME) between internet use and mental health outcomes. Grey rectangles indicate missing effect sizes because the model failed to converge with the limited sample size of subpopulations in China and Poland, as well as the genetic data is not available in all countries except for England and USA. AME = average marginal effect, ADL = activities of daily living, PGS = polygenic scores, DS = depressive symptoms, SWB = subjective well-being.



Supplementary Fig.9: Associations of Internet use frequency with self-reported health. Data are presented as the AME with 95% CIs. The AME for each country was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.



Supplementary Fig.10: Associations of cumulative Internet use with self-reported health. Data are presented as the AME with 95% CIs. The AME for each country was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall AME (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.

References:

Miilunpalo, S., Vuori, I., Oja, P., Pasanen, M., & Urponen, H. (1997). Self-rated health status as a health measure: The predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. *Journal of Clinical Epidemiology*, *50*(5), 517–528. https://doi.org/10.1016/s0895-4356(97)00045-0

Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., Chisholm, D., Collins, P. Y., Cooper, J. L., Eaton, J., Herrman, H., Herzallah, M. M., Huang, Y., Jordans, M. J. D., Kleinman, A., Medina-Mora, M. E., Morgan, E., Niaz, U., Omigbodun, O., … UnÜtzer, Jü. (2018). The Lancet Commission on global mental health and sustainable development. *Lancet*, *392*(10157), 1553–1598. https://doi.org/10.1016/S0140-6736(18)31612-X

Peleg, S., & Nudelman, G. (2021). Associations between self-rated health and depressive symptoms among older adults: Does age matter? *Social Science & Medicine*, *280*, 114024. https://doi.org/10.1016/j.socscimed.2021.114024

World Health Organization. (2022). *Mental health*. https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response

COMMENT 1-3. *A related issue is the lack of motivating statements regarding policy implications of estimating heterogeneous effects, particularly by genetic risks. It is certainly true that effects may vary by sociodemographic and genetic characteristics. Why is it so important to look at such heterogeneity? Does it improve the way we promote mental health in a meaningful manner? In clinical settings where "precision-medicine" interventions are feasible, it may be useful to have a better understanding of effect heterogeneity. The authors briefly mentioned the importance of tailored digital inclusion. What is it? Is it feasible to realistically implement in policy settings? I think the manuscript overall lacks clarity on how the findings can be useful from policy/clinical perspectives.*

RESPONSE 1-3. Thank you very much for the comments. We agree with the reviewer on the need for clarity on the policy implications of our findings. Considering the limited understanding of the Internet's influence on mental health among older people and the potential different role of Internet use across countries, subpopulations, and individuals, such heterogeneity observed in our study is much needed for facilitating targeted Internet use-based intervention strategies and precision mental health. Our results indicate that mental health promotion initiatives should be tailored to subpopulations and even individuals based on their genetic, sociodemographic, and behavioral characteristics, which requires randomized controlled trials to provide further evidence. Following the reviewer's suggestion, we added the motivations to assess whether the relationship between Internet use and mental health differs by genetic, sociodemographic, and behavioral factors in the

Introduction Section. Additionally, we extended our discussion on the policy implications of our results regarding the heterogeneous associations between Internet use and mental health among subpopulations. In the **Conclusion** Section, "**Tailored digital inclusion strategies**" refers to tailoring digital inclusion strategies to individual needs and characteristics (Asmar et al., 2022; Lu et al., 2022), which focuses on enabling people to use the Internet properly with acquired digital skills, so that they can access the Internet and other Internet-based services. Our study revealed the protective association between Internet use and mental health with heterogeneity, highlighting the importance of not only digital inclusion but also proper and tailored digital inclusion. Further studies are needed to provide more robust evidence.

The added parts are copied below:

Introduction Section, **Third** Paragraph: "Additionally, genetic and environmental factors may interact with each other and affect the risk of mental disorders such as depression[31]. There is evidence that the association between media use/Internet use and mental health was attributed to genetic vulnerabilities in young adults[32,33]. However, few studies have explored the interaction between genetic risk and Internet use and its influence on mental health in older populations. Evaluating how genetic, sociodemographic, and behavioral factors influence the relationship between Internet use and mental health is crucial for pinpointing which subpopulations would most benefit from interventions involving Internet use[34]. Recognizing the varied impact of Internet use is important for precision mental health promotion, which currently faces challenges in systematically incorporating genetic, sociodemographic, and behavioral characteristics into the selection of mental health intervention strategies[34]."

Discussion Section, **Eighth** Paragraph: "Overall, the observed heterogeneity in our study indicates the need for precision mental health promotion that matches people to the interventions based on their genetic, sociodemographic, and behavioral factors. Take an example, using mobile health apps to encourage fitness and monitor wellness may be more beneficial for mental health in older adults without sufficient physical activity, while the interventions can be other formats, such as e-learning or online hobbies, for those who exercise regularly (exercise information can be obtained from these aging cohort surveys; i.e. relates to physical inactivity factor in our study)[57]. In the future, Internet-based mental health interventions could even be tailored to individual needs by developing different digital tools[58]. However, evaluating the effectiveness of Internet use in subpopulations of older adults may be challenging, and more robust evidence from randomized controlled trials is required."

References:

Asmar, A., Mariën, I., & Van Audenhove, L. (2022). No one-size-fits-all! Eight profiles of digital inequalities for customized inclusion strategies. *New Media & Society*, *24*(2), 279–310. https://doi.org/10.1177/14614448211063182

Lu, X., Yao, Y., & Jin, Y. (2022). Digital exclusion and functional dependence in older people: Findings from five longitudinal cohort studies. *eClinicalMedicine*, *54*, 101708. https://doi.org/10.1016/j.eclinm.2022.101708

COMMENT 1-4. *For the main analysis, many aspects of their analytic approach remained unclear after reading what's written in the manuscript. The authors wrote that they used multiple waves of data and modeled it with mixed-effects regression with years since baseline as the timescale. This suggests that they specified a functional form (presumably linear) for the time trend in the outcome and added the internet use variable as a separate term. This model effectively assumes that the association between internet use and outcomes does not vary over time. This is an assumption that is hard to believe and in fact contradicts their own empirical findings about the effect heterogeneity by age. Providing an equation for the regression model would be appreciated.*

RESPONSE 1-4. Thank you so much for the valuable comments and suggestions. We agree that assuming the association between Internet use and outcomes remains constant over time is a strong assumption. Following the reviewer's suggestion, we included the interaction between Internet use and follow-up time in the linear mixed model, and calculated the average marginal effect as the effect size. The equations for the linear mixed model and average marginal effect are provided in the **Methods** Section. From the predicted standardized scores for three health indicators (**Supplementary Fig.2–4**), we observed that the protective association between Internet use and mental health remained unchanged over time, with varying effect sizes in some countries (e.g., Brazil, China, and Poland). Thus, the revised results were consistent with the findings without considering the interaction term. We added the following descriptions to the **Methods** Section and modified the corresponding results.

Methods Section, **Statistical analyses** Part: "The basic model included Internet use, years since baseline, and their interaction. All models were then adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The linear mixed model can be expressed as equation (1):

$$y_{ij} = \beta_0 + \beta_1 Internet_i + \beta_2 Time_{ij} + \beta_3 Internet_i * Time_{ij} + X_i\beta' + b_{0i} + \epsilon_{ij} \quad (1)$$
$$i = 1, \dots, N; j = 1, \dots, n_i$$

where y_{ij} is the mental health outcome for the i-th individual measured at the j-th follow-up; $Internet_i$ denotes the Internet use information at baseline; $Time_{ij}$ is the year since baseline; X_i is the $N \times p$ design matrix for baseline covariates with β' corresponding a $p \times 1$ vector of coefficients; $b_{0i} \sim N(0, \sigma_b^2)$ is the random effect for i-th individual; and $\epsilon_{ij} \sim N(0, \sigma^2)$ is the random error. We calculated the average marginal effect (AME) using equation (2) to illustrate the mean change in the estimated values of each standardized outcome from not using the Internet to using the Internet (marginaleffects package)[87].

$$\mathbb{E}(y_{ij} | Internet_i = 1) - \mathbb{E}(y_{ij} | Internet_i = 0) \quad (2)$$

The AME is an informative way to evaluate the treatment effect especially in the nonlinear regression model[88]."

Results Section, **Associations between baseline Internet use and mental health** Part: "Longitudinal associations between baseline Internet use and subsequent mental health are shown in Fig.1 and Supplementary Fig.1. After adjusting for demographic characteristics, socioeconomic status, health behaviors, and physical health, Internet use was associated with fewer depressive symptoms (pooled average marginal effect [AME] = -0.09, 95% confidence interval [CI] = -0.12 to -0.07, I^2 = 70.68%, H^2 = 3.41), higher life satisfaction (pooled AME = 0.07, 95% CI = 0.05 to 0.10, I^2 = 72.68%, H^2 = 3.66), and better self-reported health (pooled AME = 0.15, 95% CI = 0.12 to 0.17, I^2 = 73.53%, H^2 = 3.78). Predicted standardized scores for depressive symptoms, life satisfaction, and self-reported health showed that the protective association between Internet use and mental health remained unchanged over time, with varying effect sizes in some countries, such as Brazil, China, and Poland (Supplementary Fig.2–4).”

The updated figures are copied below:

In the main text:

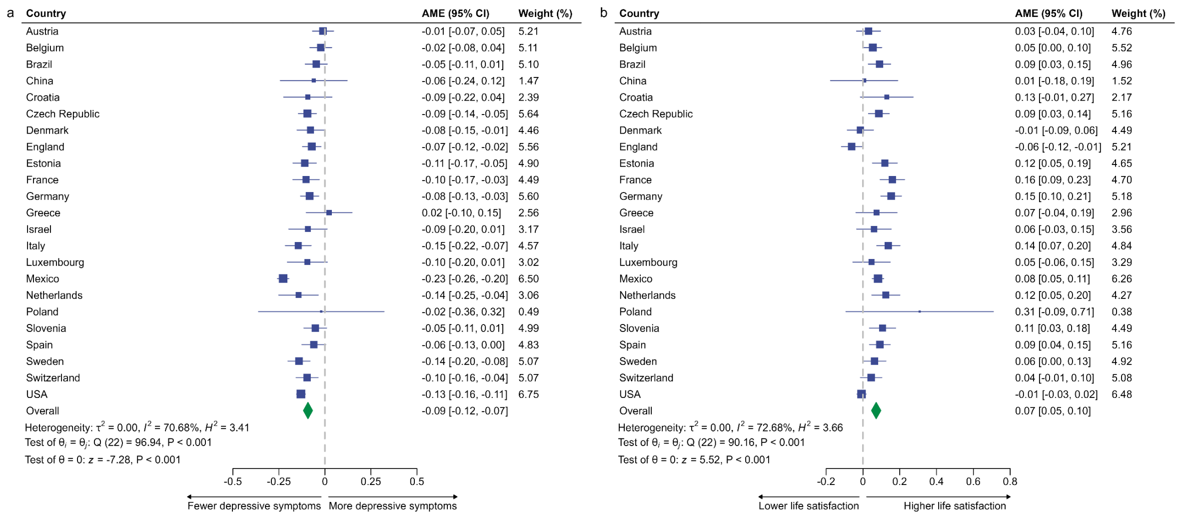
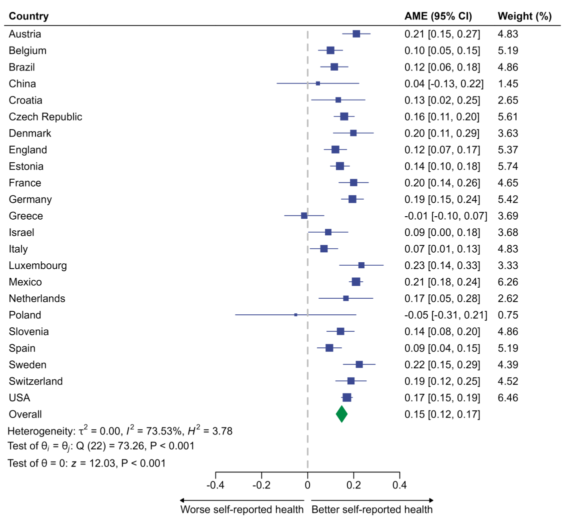
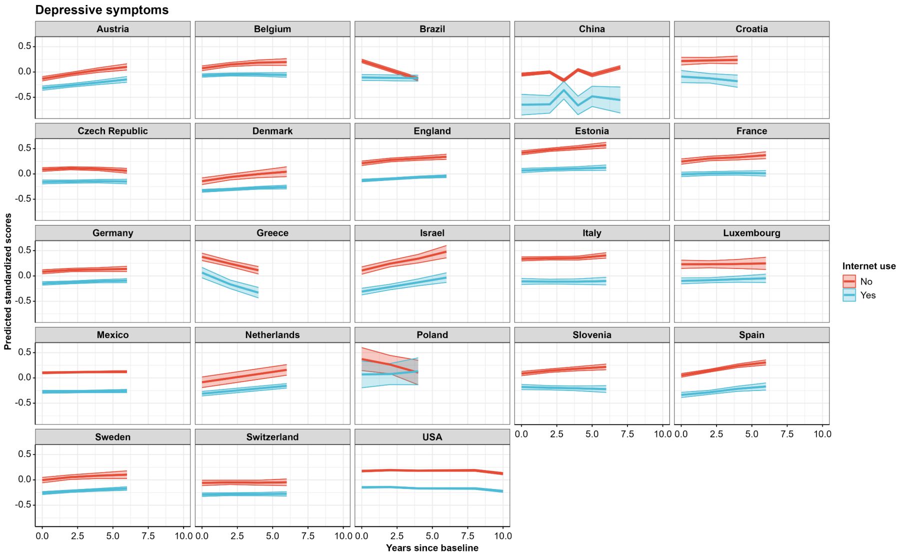


Fig.1: Associations of Internet use with depressive symptoms (a) and life satisfaction (b). Data are presented as the AME with 95% CIs. The AME for each country (blue square) was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall AME (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.

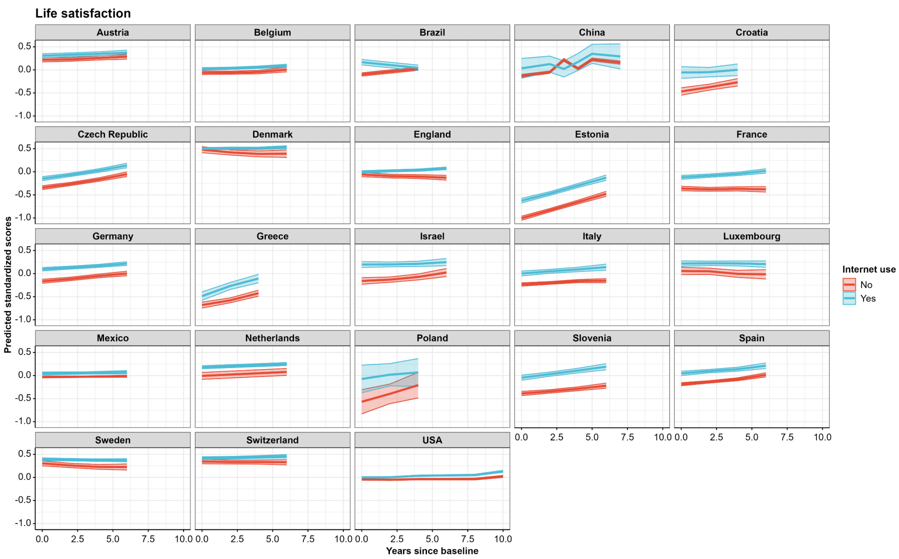
In the Supplementary Materials:



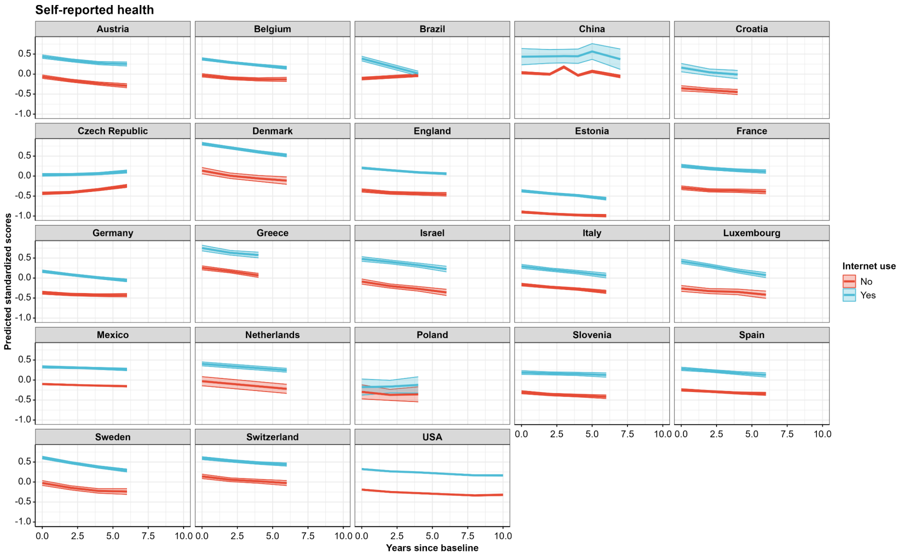
Supplementary Fig.1: Associations of Internet use with self-reported health. Data are presented as the AME coefficients with 95% CIs. The AME for each country (blue square) was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall AME (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.



Supplementary Fig.2: Predicted standardized scores for depressive symptoms over time by baseline Internet use. Data are presented as the predictions with 95% CIs. The prediction was calculated using the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL, shown as equation (2). CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.



Supplementary Fig.3: Predicted standardized scores for life satisfaction over time by baseline Internet use. Data are presented as the predictions with 95% CIs. The prediction was calculated using the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL, shown as equation (2). CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.



Supplementary Fig.4: Predicted standardized scores for self-reported health over time by baseline Internet use. Data are presented as the predictions with 95% CIs. The prediction was calculated using the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL, shown as equation (2). CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.

COMMENT 1-5. *It was also unclear why the authors provided the joint effect analysis only for genes. Unless there are ways to intervene and modify genes, the interaction analysis should suffice for genes as well. This way, the authors can combine the results for all other sociodemographic characteristics with those for genes.*

RESPONSE 1-5. Thank you very much for this helpful suggestion. We agree with the reviewer that analyzing the interaction between Internet use and genetic risks is sufficient in our study because modifying genes is unrealistic in current practice. Therefore, we removed the results regarding the joint effect analyses. We also updated **Fig.2** and **Supplementary Fig.8** by adding the interaction analyses for genetic risks. We copied the updated results below:

Results Section, **Subpopulation analyses** Part: "We assessed genetic risk levels by polygenic scores for depressive symptoms and subjective well-being in England and the USA (Fig.2 and Supplementary Fig.8). After stratifying participants by the genetic risk (i.e., low [quintile 1], intermediate [quintile 2–4], and high [quintile 5]) of depression, Internet use was beneficially associated with depressive symptoms and self-reported health across three genetic risk categories in England and the USA. However, we observed that the association between Internet use and life satisfaction was significant only among participants with intermediate genetic risk of depression in England. The results of replacing polygenic scores for depressive symptoms with polygenic scores for subjective well-being were similar. We only observed significant interaction between Internet use and genetic risk of depression on life satisfaction in England (P = 0.032)."

The modified figure is copied below:

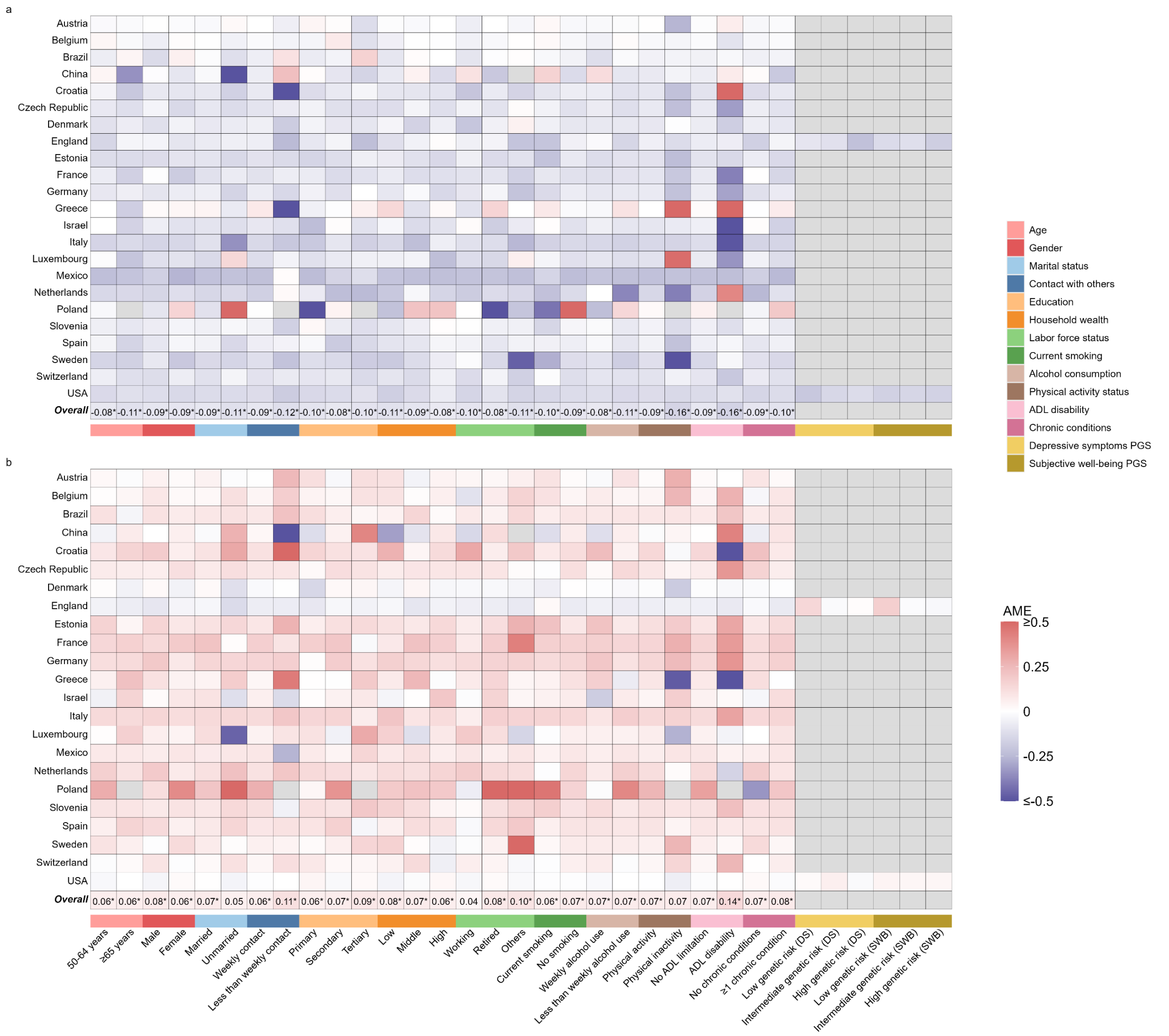
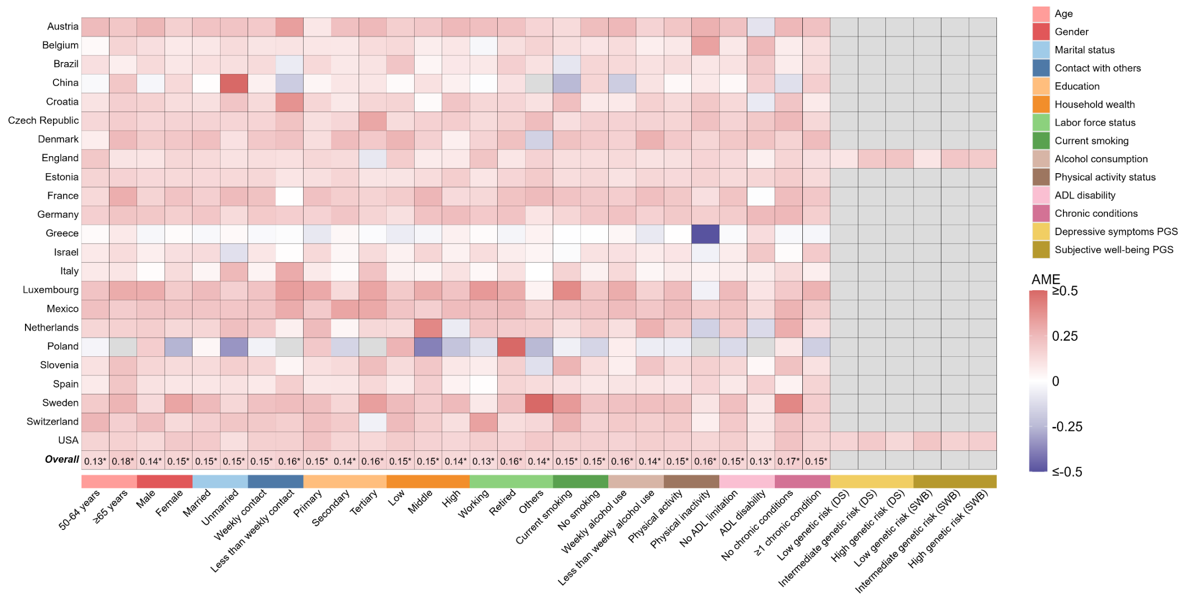


Fig.2: Associations of Internet use with depressive symptoms (a) and life satisfaction (b) by age, gender, marital status, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and genetic risks assessed by PGS for depressive symptoms and subjective well-being. The AME for each country was separately estimated by the linear mixed model adjusted for all

covariates (not including genetic risks) except the stratification variable. The overall AME was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. The color of cells indicates the effect sizes (AME) between internet use and mental health outcomes. Grey rectangles indicate missing effect sizes because the model failed to converge with the limited sample size of subpopulations in China and Poland, as well as the genetic data is not available in all countries except for England and USA. AME = average marginal effect, ADL = activities of daily living, PGS = polygenic scores, DS = depressive symptoms, SWB = subjective well-being.



Supplementary Fig.8: Associations of Internet use with self-reported health by age, gender, marital status, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and genetic risks assessed by PGS for depressive symptoms and subjective well-being. The AME for each country was separately estimated by the linear mixed model adjusted for all covariates (not including genetic risks) except the stratification variable. The overall AME was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. The color of cells indicates the effect sizes (AME) between internet use and mental health outcomes. Grey rectangles indicate missing effect sizes because the model failed to converge with the limited sample size of subpopulations in China and Poland, as well as the genetic data is not available in all countries except for England and USA. AME = average marginal effect, ADL = activities of daily living, PGS = polygenic scores, DS = depressive symptoms, SWB = subjective well-being.

COMMENT 1-6. *The authors analyzed those who responded to all waves. This means that the sample selection occurred after the baseline treatment (i.e., the internet use), leading to selection bias (also known as collider stratification bias). They further removed individuals with missing data, further increasing the risk of selection bias. For the former point, the authors are encouraged to compare baseline characteristics between the whole baseline participants (e.g., wave 9 participants in HRS) and the analytic samples. If the attrition is* *associated with the treatment status and other covariates that are potentially associated with the outcomes, then selection bias is going to be a concern.*

RESPONSE 1-6. We thank the reviewer for the insightful comments. To clarify our exclusion criteria, we followed a common practice in cohort study by excluding participants with missing information on Internet use and three outcomes at baseline, as well as those without the reassessment of three outcomes at any follow-up surveys (**Methods** Section, **Study population** Part). We did not exclude participants who did not have data in all waves. We agree with the reviewer that the selection bias, which is common in the observational studies (Sedgwick, 2014), may be a concern in our study. We observed differences in baseline characteristics between the analytic samples and the entire populations from the original dataset in six cohorts (please refer to **Supplementary Table 5**, which is too long to be presented here). To address this issue, we created the inverse probability weights of attrition and fitted linear mixed models with these weights to account for selection bias due to attrition. The updated results were consistent with our main findings. We also added the discussion of the selection bias as a limitation in the **Discussion** section. The added descriptions are provided below:

In the main text:

Methods Section, **Statistical analyses** Part: "Third, to account for selection bias due to attrition, we created the inverse probability weights of attrition and fitted linear mixed models with these weights. Additionally, we compared the baseline characteristics between the analytic samples for the Internet use and mental health study and the entire populations from the original dataset in six cohorts."

Results Section, **Sensitivity analyses** Part: "After excluding participants from England and Mexico (Supplementary Fig.7) or using the inverse probability weights of attrition in the linear mixed model (Supplementary Fig.15–17), we observed that the results were consistent with our main findings."

Discussion Section, **Twelfth** Paragraph: "Second, selection bias may exist in our study since there were differences in baseline characteristics between the analytic samples and the entire populations from the original dataset in six cohorts (Supplementary Table 5), even though the results using the inverse probability weights of attrition were similar. Larger cohort studies are needed to further verify the results."

In the Supplementary Materials:

Supplementary Methods Section, **Inverse probability weights of attrition** Part: "To account for selection bias due to attrition, we created the stabilized inverse probability weights of attrition using the following equation[1]:

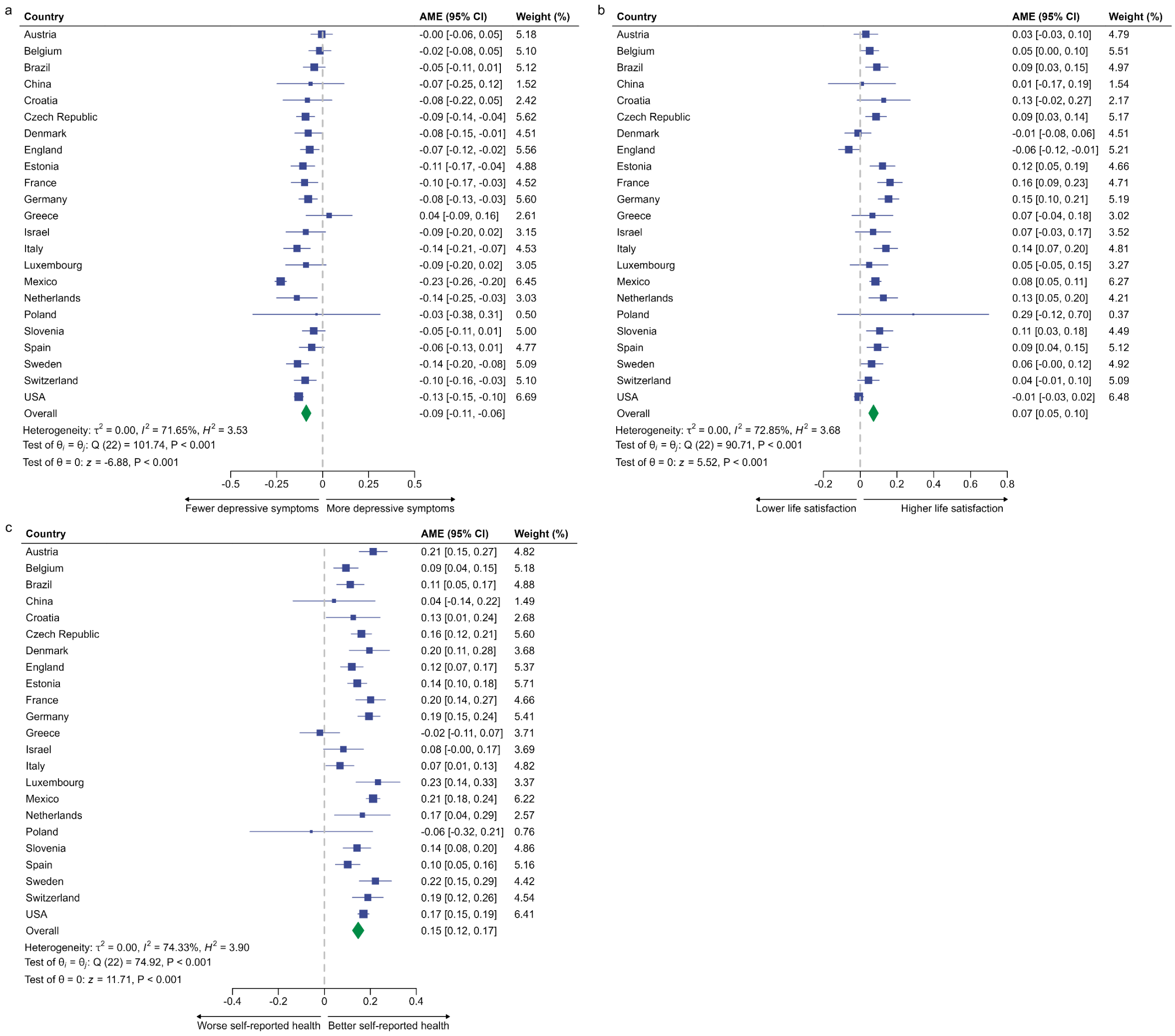
$$IPW_{ij} = \prod_{k=0}^{j} \frac{\Pr\left\{C_{ik} = 0 \middle| C_{i(k-1)} = 0, X_i, V_i\right\}}{\Pr\left\{C_{ik} = 0 \middle| C_{i(k-1)} = 0, X_i, \bar{L}_{i(k-1)}\right\}}$$

where $C_{ik} = 0$ indicates that participant i is not censored by wave k; X_i represents the baseline Internet use; V_i represents fourteen baseline covariates in our study; $\bar{L}_{i(k-1)}$ represents time-varying variables at the previous wave (i.e., wave $k-1$), including standardized values for depressive symptoms, life satisfaction, and self-reported health. The numerator represents the conditional probability of remaining uncensored up to wave k, given the baseline covariates and Internet use. Similarly, the denominator represents the conditional probability of remaining uncensored up to wave k, given the baseline covariates, Internet use, and standardized scores for three health outcomes at the previous wave. We estimated these probabilities using pooled logistic regression models. The weights for each participant at each wave were proportional to the inverse of the wave-specific probability of remaining uncensored. Since there were only two waves of follow-up in Brazil, we estimated the inverse probability weights of attrition as follows:

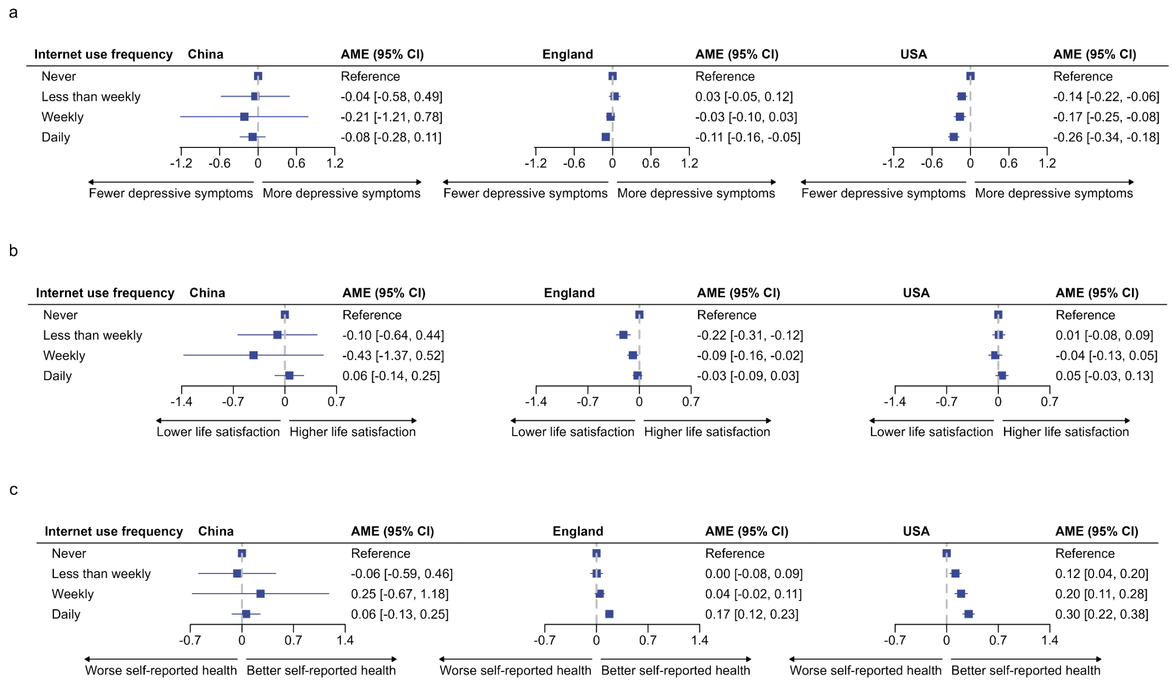
$$IPW_i = \frac{1}{\Pr\{C_i = 0 | X_i, V_i\}}$$

".

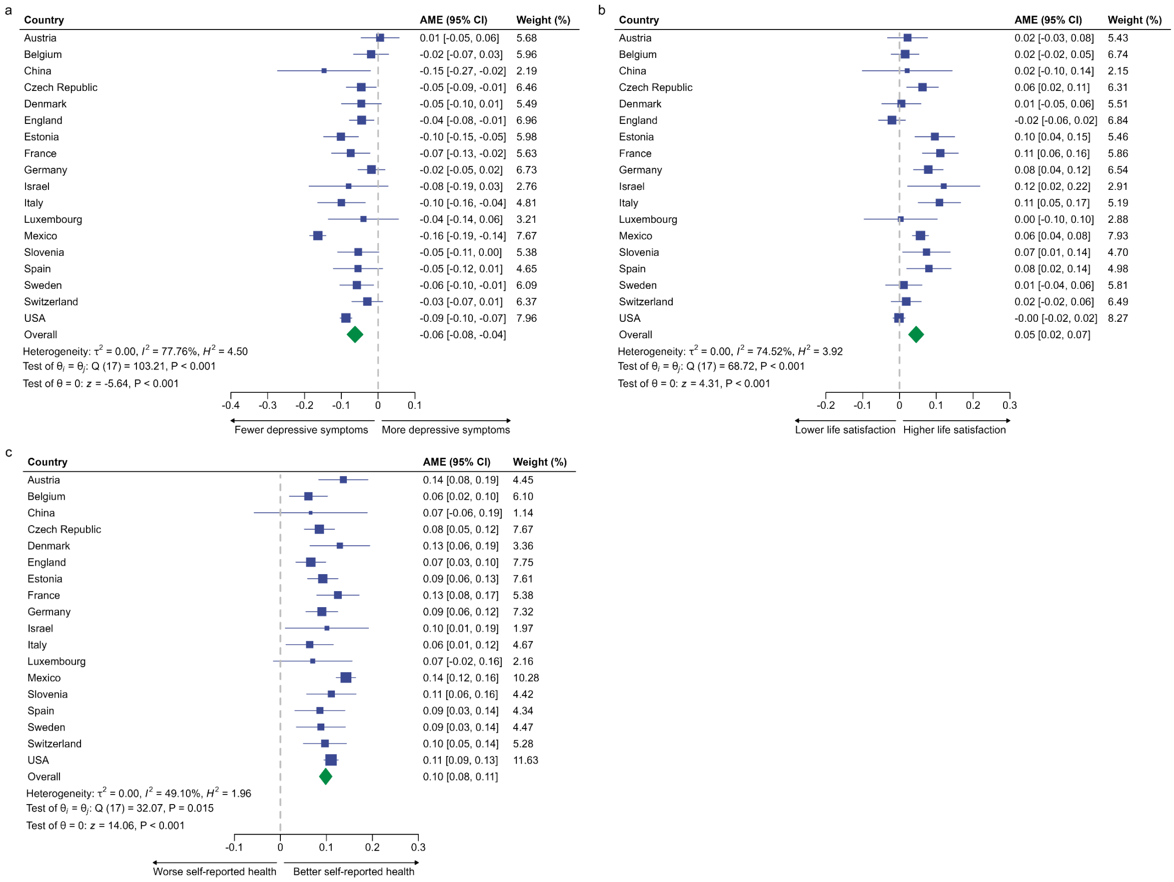
The added figures and tables are copied below:



Supplementary Fig.15: Associations of Internet use with depressive symptoms (a), life satisfaction (b), and self-reported health (c) using the inverse probability weights of attrition. Data are presented as the AME with 95% CIs. The AME for each country (blue square) was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall AME (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.



Supplementary Fig.16: Associations of Internet use frequency with depressive symptoms (a), life satisfaction (b), and self-reported health (c) using the inverse probability weights of attrition. Data are presented as the AME with 95% CIs. The AME for each country (blue square) was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.



Supplementary Fig.17: Associations of cumulative Internet use with depressive symptoms (a), life satisfaction (b), and self-reported health (c) using the inverse probability weights of attrition. Data are presented as the AME with 95% CIs. The AME for each country (blue square) was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall AME (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.

References:

Sedgwick, P. (2014). Bias in observational study designs: Prospective cohort studies. *BMJ (Clinical Research Ed.)*, *349*, g7731. https://doi.org/10.1136/bmj.g7731

COMMENT 1-7. *For the second point, the authors showed results based on multiple imputation. I think it makes more sense to show these findings as the main results rather than showing them as sensitivity analyses.*

RESPONSE 1-7. Thank you for the suggestion. Following the reviewer's suggestion, we replaced **Fig.1–5** in the revised manuscript with the results using imputed datasets. We also moved the results based on complete cases in the Supplementary Materials as sensitivity analyses (**Supplementary Fig.12–14**). The added part and updated figures are copied below:

Results Section, **Sensitivity analyses** Part: "We re-ran the main analyses using complete cases, and the results were largely replicated (Supplementary Fig.12–14)."

In the main text:

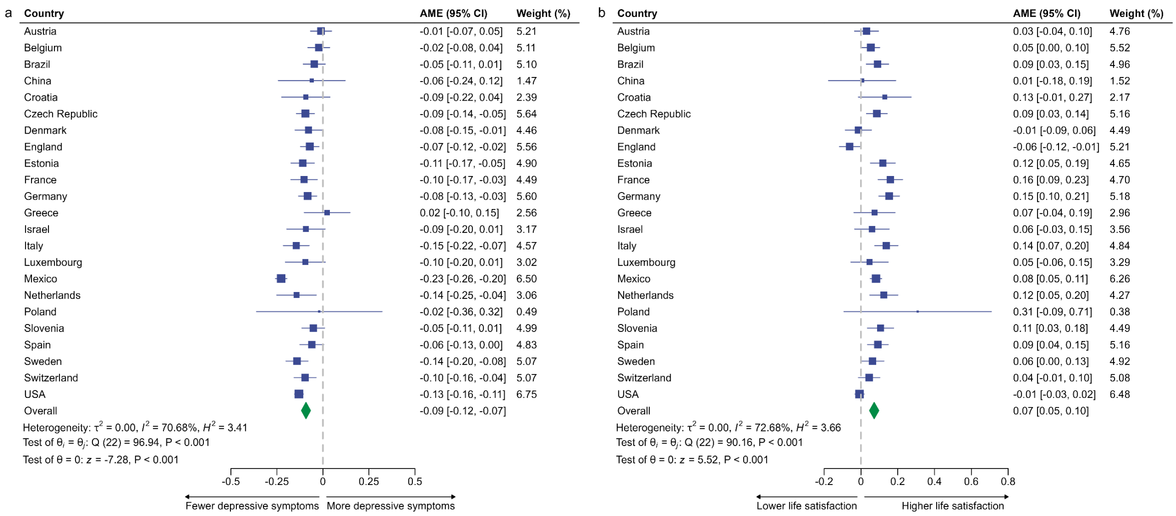


Fig.1: Associations of Internet use with depressive symptoms (a) and life satisfaction (b). Data are presented as the AME with 95% CIs. The AME for each country (blue square) was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall AME (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.

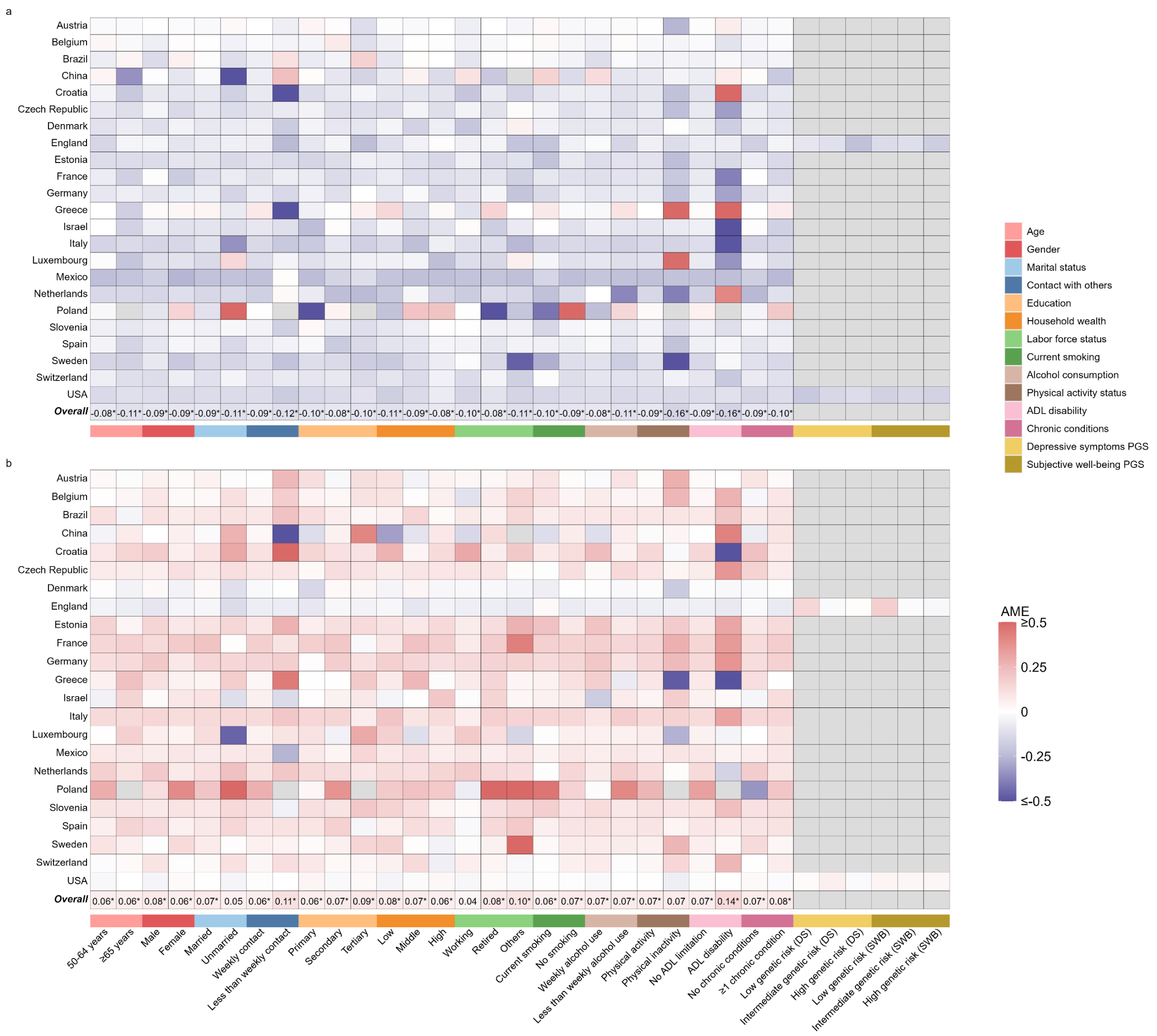


Fig.2: Associations of Internet use with depressive symptoms (a) and life satisfaction (b) by age, gender, marital status, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and genetic risks assessed by PGS for depressive symptoms and subjective well-being. The AME for each country was separately estimated by the linear mixed model adjusted for all covariates (not including genetic risks) except the stratification variable. The overall AME was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. The color of cells indicates the effect sizes (AME) between internet use and mental health outcomes. Grey rectangles indicate missing effect sizes because the model failed to converge with the limited sample size of subpopulations in China and Poland, as well as the genetic data is not available in all countries except for England and USA. AME = average marginal effect, ADL = activities of daily living, PGS = polygenic scores, DS = depressive symptoms, SWB = subjective well-being.

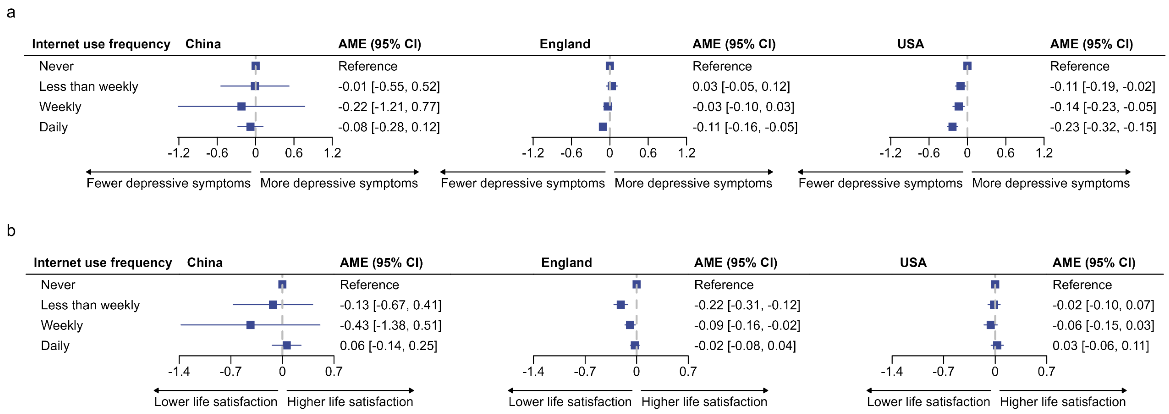


Fig.3: Associations of Internet use frequency with depressive symptoms (a) and life satisfaction (b). Data are presented as the AME with 95% CIs. The AME for each country was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.

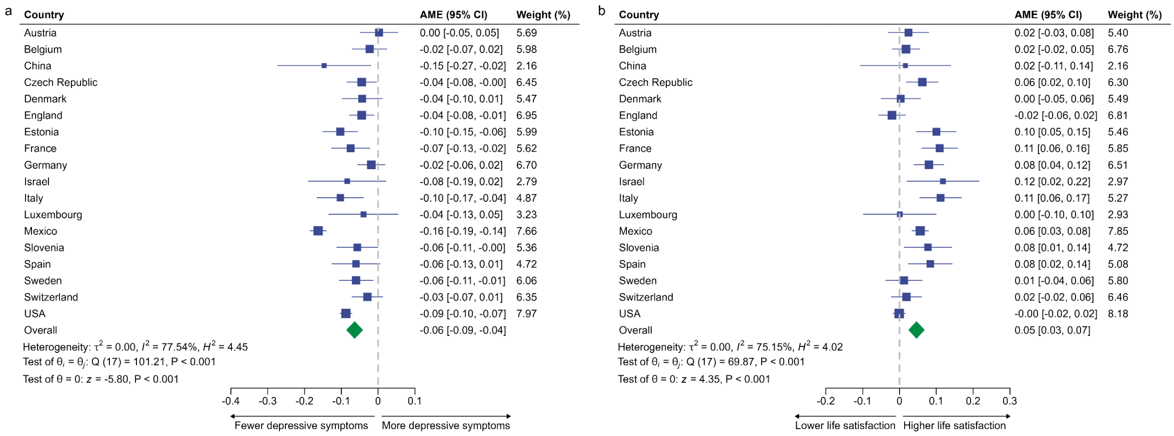


Fig.4: Associations of cumulative Internet use with depressive symptoms (a) and life satisfaction (b). Data are presented as the AME with 95% CIs. The AME for each country (blue square) was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall AME (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.

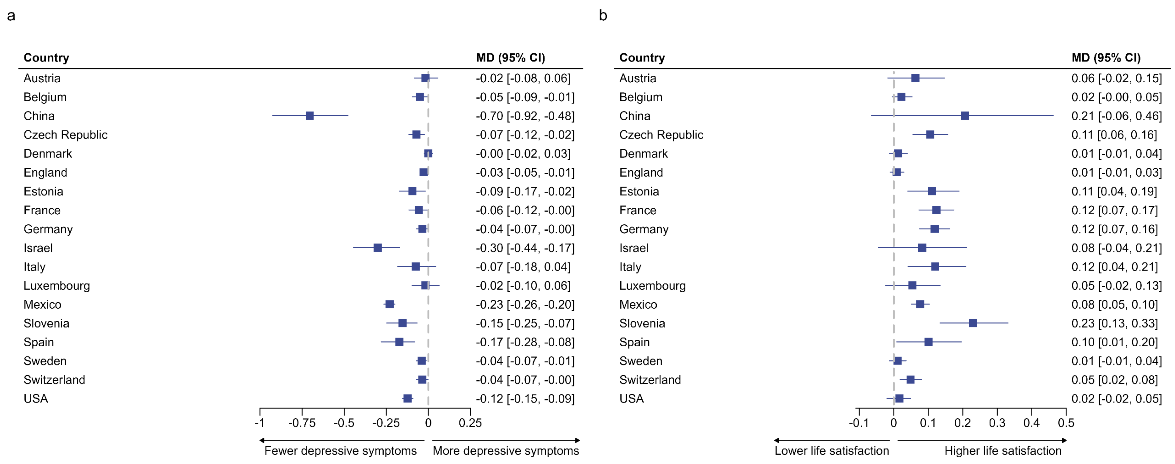
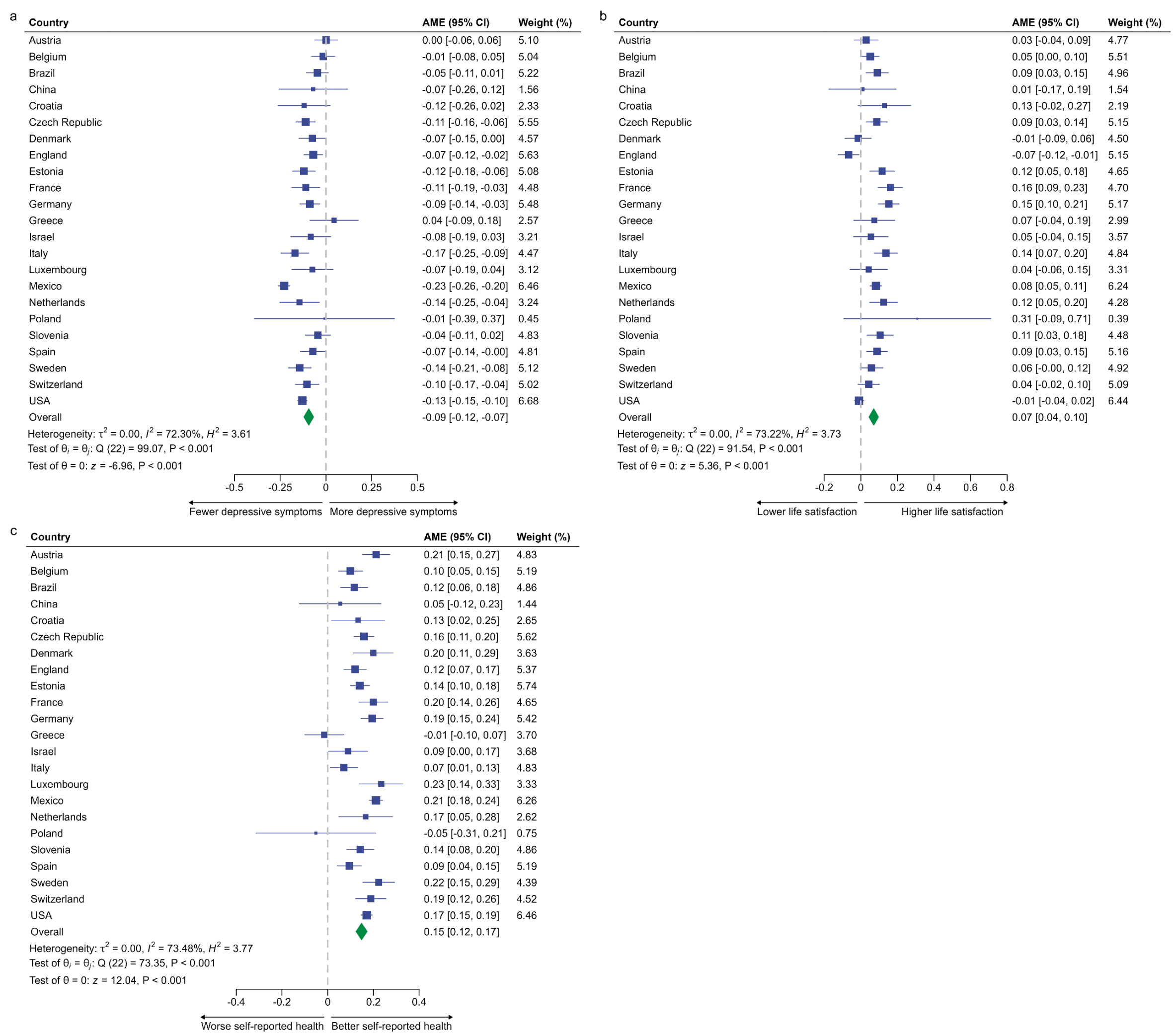
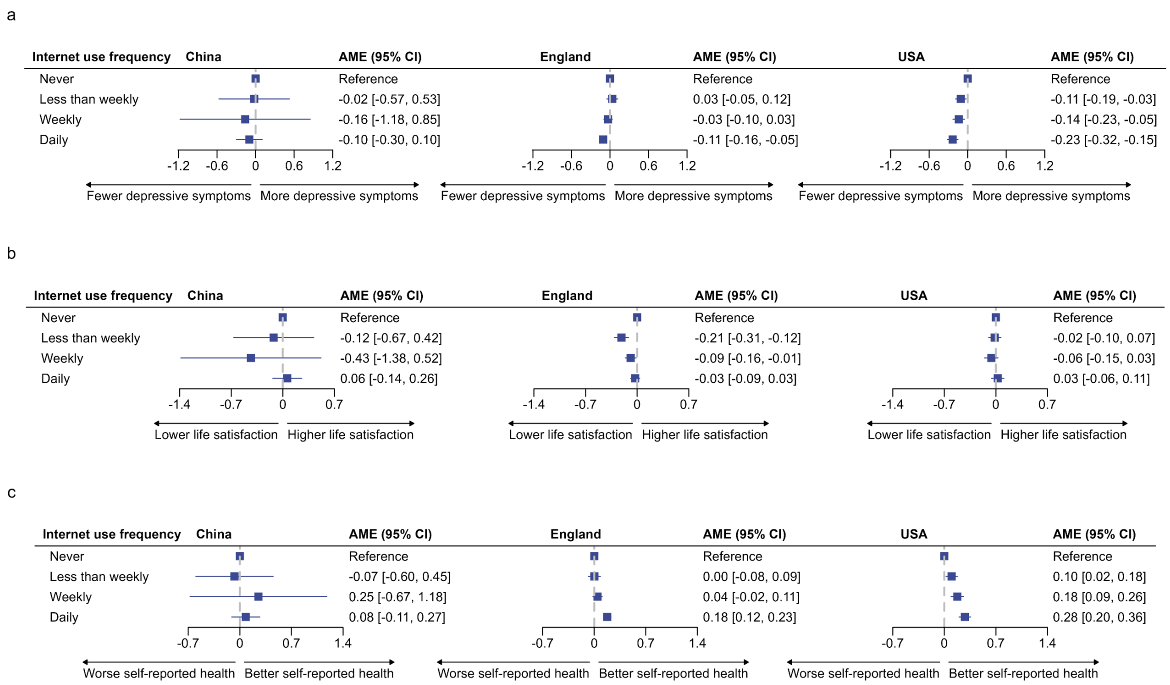


Fig.5: Estimated mean difference in the standardized scores of depressive symptoms (a) and life satisfaction (b) in the last follow-up between populations following the natural course (as the reference group) and those who received the intervention consistently. Data are presented as the MDs with 95% CIs. MD = mean difference, CI = confidence intervals.

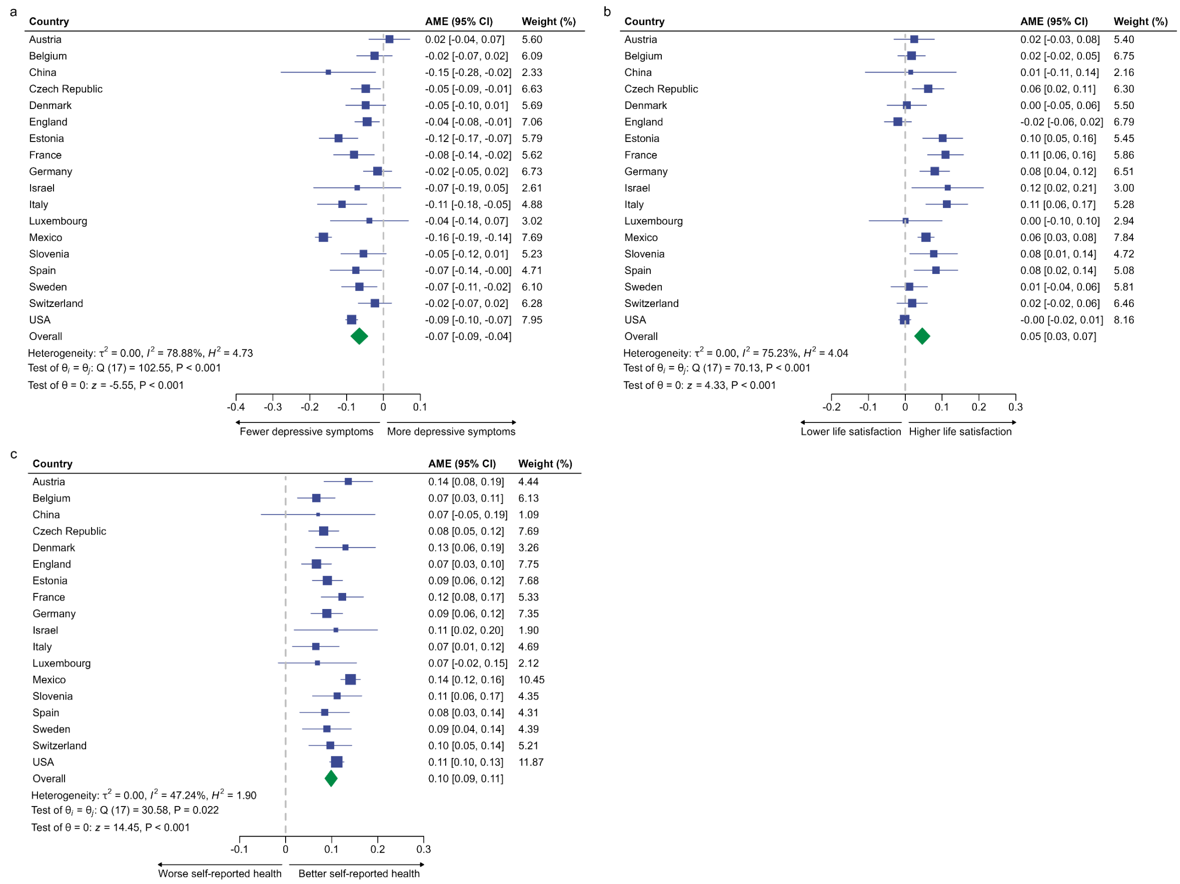
In the Supplementary Materials:



Supplementary Fig.12: Associations of Internet use with depressive symptoms (a), life satisfaction (b), and self-reported health (c) using complete cases. Data are presented as the AME with 95% CIs. The AME for each country (blue square) was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall AME (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.



Supplementary Fig.13: Associations of Internet use frequency with depressive symptoms (a), life satisfaction (b), and self-reported health (c) using complete cases. Data are presented as the AME with 95% CIs. The AME for each country was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.



Supplementary Fig.14: Associations of cumulative Internet use with depressive symptoms (a), life satisfaction (b), and self-reported health (c) using complete cases. Data are presented as the AME with 95% CIs. The AME for each country (blue square) was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall AME (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.

COMMENT 1-8. *Their analysis of cumulative internet use is essentially flawed because it ignores the presence of time-dependent confounders. It is well known that the estimation of a sustained treatment (e.g., cumulative internet use) requires more sophisticated approaches for confounding adjustment (e.g., g-computation).*

RESPONSE 1-8. We thank the reviewer for the insightful comments. Here, we aim to characterize the association between cumulative Internet usage and mental health. Therefore, we adopted a common approach to quantify the cumulative Internet use as the number of waves where participants used the Internet during the first two biennial waves (Cho et al., 2023; Yu, Cho, et al., 2023; Yu, Westrick, et al., 2023). We agree with the reviewer that more sophisticated approaches are needed to further validate the results by accounting for the time-varying confounders. Following the reviewer's suggestion, we further used the parametric g-formula to estimate the effect of Internet use on depressive symptoms, life satisfaction, and self-reported health at the end of the follow-up period. The results were consistent with our main findings. This further analysis also provides solid evidence of the beneficial role of Internet use, which is consistent with the association between baseline Internet use and mental health (**Fig.1**). The results of g-formula are presented as **Fig.5** in the revised manuscript. Detailed methods, results, and discussions are copied below:

In the main text:

Methods Section, **Statistical analyses** Part: "To further consider the influence of time-varying confounders, we used the parametric g-formula to examine the effect of sustained Internet use on mental health outcomes at the end of the follow-up period (gfoRmula package). The parametric g-formula is a statistical method to estimate the causal effect of sustained treatment strategies and appropriately adjust for measured time-varying confounders[91]. Detailed algorithms are provided in Supplementary Methods."

Results Section, **Associations of frequency and cumulative period of Internet use with mental health** Part: "To adjust for time-varying confounders while modeling dynamic interventions (i.e., Internet use), we used the parametric g-formula to estimate the effect of sustained Internet use on mental health outcomes (Fig.5 and Supplementary Fig.11). Results confirmed that Internet use was associated with changes in depressive symptoms (mean difference [MD] = -0.0003 in Denmark to -0.70 in China), life satisfaction (MD = 0.008 in England to 0.23 in Slovenia), and self-reported health (MD = 0.02 in England to 0.36 in China)."

In the Supplementary Materials:

Supplementary Methods Section, **Effect of cumulative Internet use on mental health** Part: "We used the parametric g-formula to estimate the effect of sustained Internet use on mental health. Here, we treated age, gender, and education as time-constant covariates, while considering other variables, including Internet use, marital status, household size, contact with others, household wealth, labor force status, current smoking, alcohol consumption, physical activity, number of chronic conditions, activities of daily living disability, number of limitations on instrumental activities of daily living, and the number of waves, as time-varying covariates. An outline of the parametric g-formula algorithm is as follows[7]:

1) Fit parametric regression models for the time-varying covariates at each wave k as a function of the intervention (i.e., Internet use) and covariate history among individuals under follow up at k (only consider the value at the wave $k - 1$).

2) Fit the linear regression models for each of the three standardized health indicators measured in the final wave (e.g., at the end of follow-up) as a function of the intervention and covariate history (consider the value at both the wave k and $k - 1$) among individuals under the final wave of follow-up.

3) Use a Monte Carlo simulation to generate populations with the same sizes of original analytic samples under each of the treatment strategies (i.e., whether to use the Internet). For each individual:

a) The values of baseline covariates at $k = 0$ are randomly sampled with replacement from the individuals in the original study population.

b) The values of the time-varying covariates at $k > 0$ are drawn from the parametric distribution estimated in Step 1.

c) The value of Internet use is set to the value specified by the strategy, e.g., one at all times k for the Internet use intervention.

d) The covariate-specific mean standardized health indicators at the end of follow-up is estimated from the model in Step 2.

4) Compute the mean standardized health indicators at the end of follow-up in the population under each strategy, and the mean differences between strategies.

5) Repeat above steps in 100 bootstrap samples in order to obtain 95% confidence intervals.

We estimated the mean difference in the standardized health indicator in the final wave between simulated populations following the natural course (as the reference group) and those who received the intervention consistently (denoted as *always-treat*). The natural course corresponds to no interventions[8]. We generated the *always-treat* populations by setting the Internet use to "yes" and simulating time-varying covariates and outcomes based on the regression models previously described[8,9]. To address the missing values, we created 10 imputed datasets using multiple imputation by chained equations. Within each imputed dataset, we bootstrapped 100 samples to derive the 95% percentile confidence interval.

Finally, we pooled the estimated mean difference and confidence interval using the subsequent equations[10]:

$$\bar{Q}_m = \frac{1}{m}\sum_{j=1}^{m}\hat{Q}_j\,;$$
$$\bar{p}_{0.025,j} = \frac{1}{m}\sum_{j=1}^{m}\hat{p}_{0.025,j}\,;$$
$$\bar{p}_{0.975,j} = \frac{1}{m}\sum_{j=1}^{m}\hat{p}_{0.975,j}$$

where $\hat{Q}_j$ indicates the point estimate from the imputed dataset j; $\hat{p}_{0.025}$ and $\hat{p}_{0.975}$ are estimates of the 2.5% and 97.5% percentiles $p_{0.025}$ and $p_{0.975}$ of the bootstrap distribution of the estimated mean differences for the imputed dataset j."

The added figures are copied below:

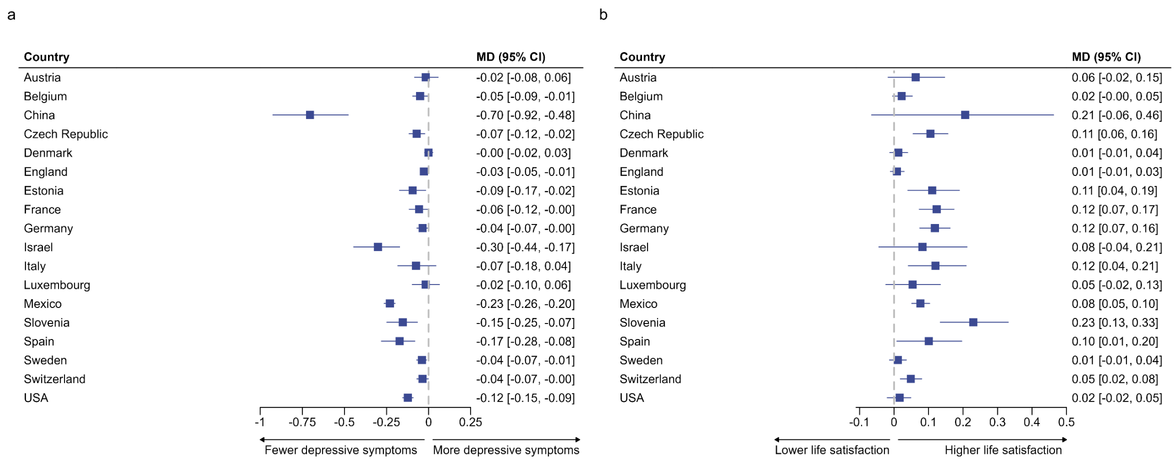
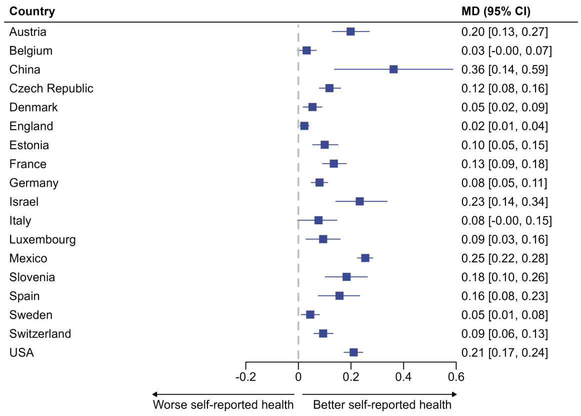


Fig.5: Estimated mean difference in the standardized scores of depressive symptoms (a) and life satisfaction (b) in the last follow-up between populations following the natural course (as the reference group) and those who received the intervention consistently. Data are presented as the MDs with 95% CIs. MD = mean difference, CI = confidence intervals.



Supplementary Fig.11: Estimated mean difference in the standardized scores of self-reported health in the last follow-up between populations following the natural course (as the reference group) and those who received the intervention consistently. Data are presented as the MDs with 95% CIs. MD = mean difference, CI = confidence intervals.

References:

Cho, G., Betensky, R. A., & Chang, V. W. (2023). Internet usage and the prospective risk of dementia: A population-based cohort study. *Journal of the American Geriatrics Society*, *71*(8), 2419–2429. https://doi.org/10.1111/jgs.18394

Yu, X., Cho, T.-C., Westrick, A. C., Chen, C., Langa, K. M., & Kobayashi, L. C. (2023). Association of cumulative loneliness with all-cause mortality among middle-aged and older adults in the United States, 1996 to 2019. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(51), e2306819120. https://doi.org/10.1073/pnas.2306819120

Yu, X., Westrick, A. C., & Kobayashi, L. C. (2023). Cumulative loneliness and subsequent memory function and rate of decline among adults aged ≥50 in the United States, 1996 to 2016. *Alzheimer's & Dementia*, *19*(2), 578–588. https://doi.org/10.1002/alz.12734

COMMENT 1-9. *Discussions of the findings should be more balanced. The discussions are generally based on the premise that the internet is good for mental well-being, and this premise is reflected in the emphasis on interpreting the meta-analyzed results and the conclusion highlighting the positive effects. The results seem to be more nuanced than that, with some results even showing adverse effects. Also, the authors' discussion of policy* *implications was mostly about China. I encourage the authors to expand that to more global contexts.*

RESPONSE 1-9. Thank you so much for the comments. We recognize the need to expand our discussion on the negative association between Internet use and mental health, as well as policy implications in the global context. Therefore, we first added a paragraph to discuss the reason for the negative association and the role of Internet use as a double-edged sword. Then, we provided further discussion on policy implications in more countries. The added parts are copied below:

Discussion Section, **Ninth and Tenth** Paragraphs: "Notably, we observed that the protective association between Internet use and mental health became non-protective in England (only for life satisfaction), and insignificant in half of other countries after adjusting for demographic characteristics, socioeconomic positions, health behaviors, and physical health status (Supplementary Fig.21–23 and Supplementary Tables 6–8). As we discussed above, the different roles of Internet use may reflect the disparities across countries, subpopulations, and even individuals. These findings highlight the need to examine the varying effect of Internet use across countries in different subpopulations within a country.

Nonetheless, Internet use may be a double-edged sword for mental health among older adults. For example, excessive online duration has found to be associated with poor mental health, possibly by interfering with other essential offline activities, including sleep, physical activity, and real-life social activity[22,23]. Additional Internet use-related risk factors for mental health include cyberbullying, online fraud, and misinformation[59–61]. However, these risk factors were not captured by our datasets. Future research should consider more details of Internet use, such as daily screen time, purposes of use, types of online activities, and the interactions with other users, to inform effective Internet use-related intervention strategies for mental health promotion[62]."

Discussion Section, **Fourth** Paragraph: "Aging in a digital world is a global trend. In the USA, England, China, Europe, and Latin America, there have been several government policies or programs by non-profit organizations to help older people access and use digital technologies for enhancing their well-being[43–47]. Our results provide encouraging evidence that Internet use can be incorporated into public health policies to improve mental health in older populations in a digital era. In particular, given the low percentage of Internet use, there is a good chance to make use of the Internet to promote mental health among older adults in low- and middle-income countries. The Internet can be a scalable platform to provide necessary services and products, such as online shopping, e-learning, social networking, and digital health, thereby affording rich opportunities to translate into better mental health in later life[48]."

Discussion Section, **Fifth and Sixth** Paragraphs: "Nevertheless, we also observed some variations in the relationship between Internet use and mental health across countries. Our meta-regression analyses revealed nonlinear associations of the effect sizes of Internet use for depressive symptoms and life satisfaction with the Gini index, HALE at birth, Internet use rate, and digital skill levels. Specifically, the improvement of mental health related to Internet use was smaller in countries with less income inequality, longer life expectancy, higher Internet use rates, or lower levels of digital skills. One possible explanation for the variation attributed to the Gini index and HALE at birth is that such countries, particularly in Europe, may already provide easy-access mental health care, which could limit the incremental benefits of Internet-based interventions[49]. Additionally, this finding may suggest that the health benefits of Internet use are determined not only by the Internet use rates but also by the level of digital skills. Even in regions with broad Internet usage such as the USA, England, and Europe, a great portion of people, especially marginalized groups (e.g., older adults), are not able to fully participate in the digital society (e.g., use digital health services) due to the lack of digital skills[50]. Besides, high Internet use rates or low digital skill levels may also carry an increased risk of problematic Internet use (e.g., Internet addiction) and telecommunication fraud, which conversely worsen mental health of older people[51,52].

Moreover, we found that the positive correlation between Internet use and self-reported health increased with the GDP per capita and world happiness index. Countries with higher GDP per capita often have better healthcare infrastructure and resources, and thereby may provide individuals with more high-quality health-related information and online health services through such as telemedicine[53]. In countries with a greater level of happiness, people might use the Internet in a more positive way, aligned with leisure activities, resulting in a favorable effect on well-being[54]."

COMMENT 1-10. *Some statements are not empirically supported. First, the authors emphasized their findings about dose-response relationships. However, looking at the figures, I do not see strong evidence of dose-response relationships except for the effect on depressive symptoms and self-reported health in the USA.*

RESPONSE 1-10. Thank you so much for the comments. We agree that the evidence for the joint dose-response effect of Internet use and genetic risk on mental health or the dose-response relationship between Internet use frequency and mental health is not strong enough. For the former part, we followed the reviewer's suggestion, removed the joint effect analysis as suggested, and deleted the corresponding descriptions (please refer to **RESPONSE 1-5**). For the latter one, we added a statement in the **Discussion** Section to clarify that our results are inconsistent across countries; more evidence is needed to investigate the association between Internet use frequency and mental health. The added statement is copied below:

Discussion Section, **Third** Paragraph: "However, the inconsistent associations between Internet use frequency and mental health especially life satisfaction suggest the need for more empirical evidence to validate our findings."

COMMENT 1-11. *Second, the conclusion emphasizes the importance of enhancing digital and health literacies. This is a stretch to say and not empirically supported by their findings.*

RESPONSE 1-11. Thank you very much for the valuable comment. We acknowledge that our findings did not provide direct empirical support for the importance of enhancing digital and health literacies. Here, the importance of enhancing digital and health literacies was indirectly derived from our stratified subpopulation analysis. Our results showed a stronger protective role of Internet use among adults with higher level of education. In the revised manuscript, we relocated this statement from the **Conclusion** Section to the **Discussion** Section as discussions of the potential practical implications this result with the help of other empirical evidence (Yang et al., 2021). The modifications are copied below:

Discussion Section, **Seventh** Paragraph: "These findings, together with another recent empirical study[56], implicate that facilitating digital inclusion for this population requires increasing access to the Internet and related devices, as well as enhancing their ability, namely digital and health literacies, to effectively use the Internet for promoting their own health and well-being[48]."

References:

Yang, B. X., Xia, L., Huang, R., Chen, P., Luo, D., Liu, Q., Kang, L. J., Zhang, Z.-J., Liu, Z., Yu, S., Li, X., & Wang, X. Q. (2021). Relationship between eHealth literacy and psychological status during COVID-19 pandemic: A survey of Chinese residents. *Journal of Nursing Management*, *29*(4), 805–812. https://doi.org/10.1111/jonm.13221

COMMENT 1-12. *The authors also mentioned the Sustainable Development Goals (SDGs) in the Conclusion without providing context in the Introduction, which I thought was strange.*

RESPONSE 1-12. Thank you for the comment. To provide the background of the SDGs, we added the following statement in the **Introduction** Section, **First** Paragraph: "With a rapidly aging population, and given the substantial unmet and unrecognized mental health needs resulting from the COVID-19 pandemic and the current cost-of-living crisis, understanding the social determinants of mental health has become essential for developing novel and cost-effective strategies to improve mental health among middle-aged and older adults worldwide[3–5]. Furthermore, this understanding aligns with the objectives of the WHO's Comprehensive Mental Health Action Plan 2013–2030 and supports the achievement of the mental health targets outlined in the Sustainable Development Goals (SDGs)[3,6,7]."

COMMENT 1-13. *I appreciate their efforts to conduct many analyses to get nuanced results and report them using a unique data visualization approach. However, the figures with tons of information are overwhelming and hard to digest for most readers, and these results are mostly not discussed anyway. I suggest the authors present more condensed results in the main text and show these granular findings as Supplement.*

RESPONSE 1-13. Thank you so much for the comment and suggestion. We acknowledge the need for the reorganization of our results to provide readers with a clearer and more digestible overview of the key findings of our study. Following the reviewer's suggestion, in the revised manuscript, we only reported the results on the associations between baseline

Internet use, frequency of Internet use, and cumulative Internet use with depressive symptoms and life satisfaction, as well as the subpopulation analyses and parametric g-formula analyses. All other results are provided in the Supplementary Materials.

COMMENT 1-14. *The analysis of country-level indicators (e.g., Gini coefficient) seems unnecessary. These are purely exploratory analyses with strong modeling assumptions. The high R-squared values do not suggest the importance of the indicator itself. Also, it was unclear why they chose to examine correlations for clearly non-linear and non-monotonic relationships.*

RESPONSE 1-14. We thank the reviewer for the comments. We agree with the reviewer that using country-level factors to explain the heterogeneity variance in the effect size across countries was an exploratory analysis. Because these analyses are important for our cross-national study for the following reasons, we moved these results to the **Supplementary Materials** (**Supplementary Fig.6**) but retained the relevant discussions in the revised manuscript. We also provided more references to justify the need to explain differences in the effect size across countries using country-level factors and the use of R-squared in the analysis (Ailshire & Carr, 2021; Borenstein et al., 2021; Rai et al., 2013).

We agree with the reviewer that the assumption regarding the linear relationship between effect sizes and country-level factors may not be fully appropriate. To address this concern, we re-performed the meta-regression incorporating the linear, quadratic, or cubic term of country-level factors terms to assess the nonlinear association between country-level factors and effect sizes of Internet use. The results were largely consistent, with a few new findings regarding nonlinear correlations of the effect sizes for depressive symptoms with the Gini index, healthy life expectancy at birth, Internet use rate, and digital skills among active population.

The revised descriptions are copied below:

Methods Section, **Measures** Country-level factors Part: "Seven country-level factors that may explain the differences in the results across countries were collected based on the literature and covering economic, social, cultural, healthcare, and Internet-related measures[4,25,72]. They were GDP per capita (2022)[73], Gini index (2018–2021)[74], world happiness index (2023)[75], Hofstede's index (most recent available)[76], HALE at birth (2019)[77], Internet use rate (2021–2022)[78], and digital skills among active population (2019)[79]."

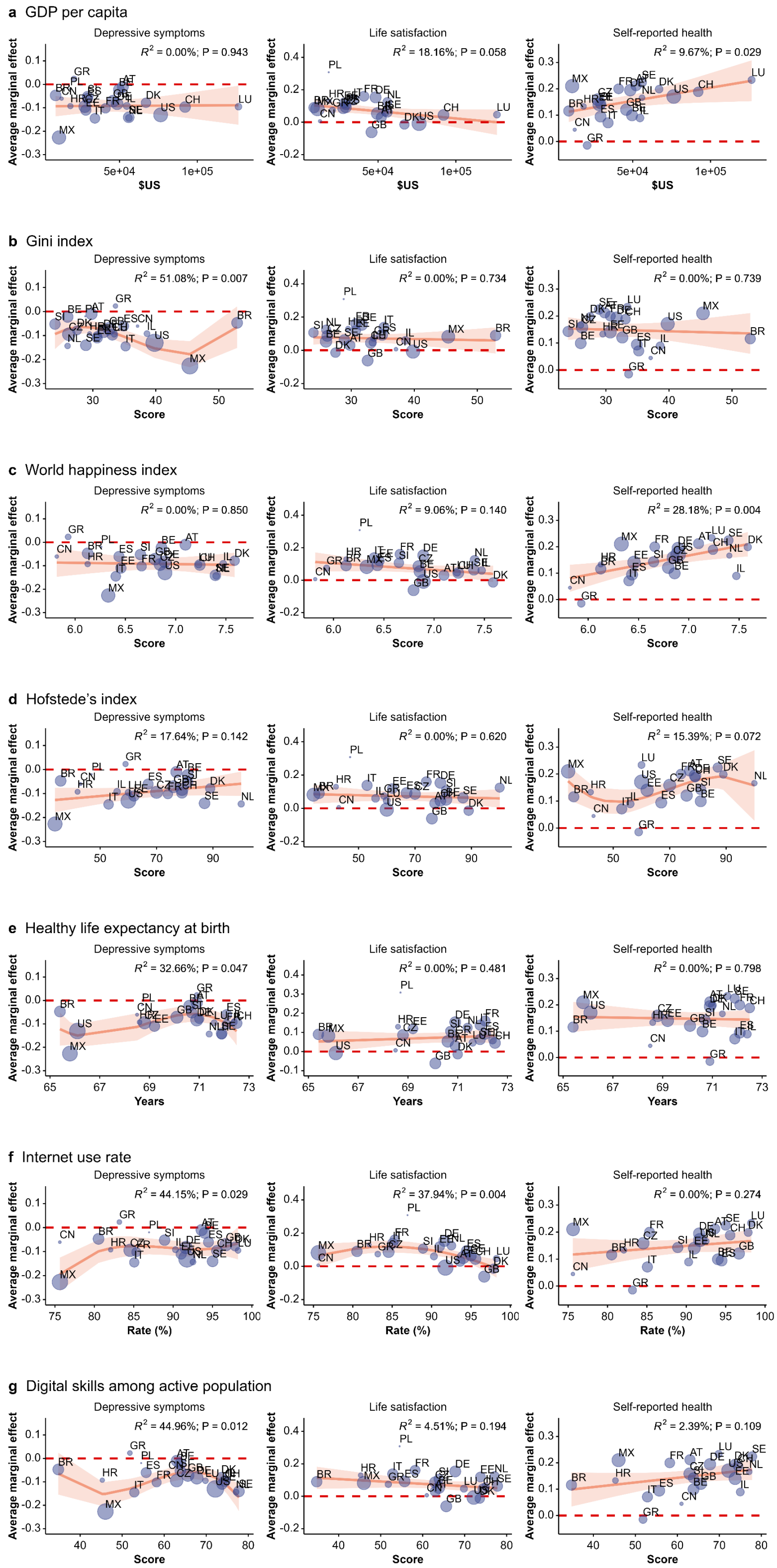
Methods Section, **Statistical analyses** Part: "To investigate the potential nonlinear relationship between effect sizes and each factor, we considered the linear, quadratic, or cubic term of country-level factors and determined the final meta-regression model if the Q-test for the model is statistically significant or if the explained variation is higher than the model with the linear term by ≥10%[89]."

Results Section, **Associations between baseline Internet use and mental health** Part: "We further explored which country-level factors may explain between-study heterogeneity in the association of Internet use with outcomes (Supplementary Fig.6). We identified nonlinear correlations of the effect sizes for depressive symptoms with the Gini index (P = 0.007), healthy life expectancy (HALE) at birth (P = 0.047), Internet use rate (P = 0.029), and digital skills among active population (P = 0.012). Moreover, nonlinear associations were observed between the effect sizes for life satisfaction and Internet use rate (P = 0.004). Regarding self-reported health, the effect sizes were linearly associated with GDP per capita (P = 0.029) and the world happiness index (P = 0.004). The correlation between the prevalence of Internet use in our samples and these country-level factors is given in Supplementary Fig.5."

Discussion Section, **Fifth and Sixth** Paragraphs: "Nevertheless, we also observed some variations in the relationship between Internet use and mental health across countries. Our meta-regression analyses revealed nonlinear associations of the effect sizes of Internet use for depressive symptoms and life satisfaction with the Gini index, HALE at birth, Internet use rate, and digital skill levels. Specifically, the improvement of mental health related to Internet use was smaller in countries with less income inequality, longer life expectancy, higher Internet use rates, or lower levels of digital skills. One possible explanation for the variation attributed to the Gini index and HALE at birth is that such countries, particularly in Europe, may already provide easy-access mental health care, which could limit the incremental benefits of Internet-based interventions[49]. Additionally, this finding may suggest that the health benefits of Internet use are determined not only by the Internet use rates but also by the level of digital skills. Even in regions with broad Internet usage such as the USA, England, and Europe, a great portion of people, especially marginalized groups (e.g., older adults), are not able to fully participate in the digital society (e.g., use digital health services) due to the lack of digital skills[50]. Besides, high Internet use rates or low digital skill levels may also carry an increased risk of problematic Internet use (e.g., Internet addiction) and telecommunication fraud, which conversely worsen mental health of older people[51,52].

Moreover, we found that the positive correlation between Internet use and self-reported health increased with the GDP per capita and world happiness index. Countries with higher GDP per capita often have better healthcare infrastructure and resources, and thereby may provide individuals with more high-quality health-related information and online health services through such as telemedicine[53]. In countries with a greater level of happiness, people might use the Internet in a more positive way, aligned with leisure activities, resulting in a favorable effect on well-being[54]."

The revised figure is copied below:



Supplementary Fig.6: Bubble plot with fitted meta regression line of the average marginal effect for three mental health outcomes and country-level factors, including GDP per capita (a), Gini index (b), world happiness index (c), Hofstede's index (d), healthy life expectancy at birth (e), Internet use rate (f), and digital skills among active population (g). Data are presented as the average marginal effects (bubbles), 95% confidence intervals (shaded area), and the linear prediction (orange line). Red dashed line at zero represents the reference line.

The R^2 quantifies the proportion of heterogeneity that can be accounted for by each country-level factor. The P value indicates whether the average marginal effect was associated with each country-level factor. AT = Austria, BE = Belgium, BR = Brazil, CN = China, HR = Croatia, CZ = Czech Republic, DK = Denmark, GB = England, EE = Estonia, FR = France, DE = Germany, GR = Greece, IL = Israel, IT = Italy, LU = Luxembourg, MX = Mexico, NL = Netherlands, PL = Poland, SI = Slovenia, ES = Spain, SE = Sweden, CH = Switzerland, and US = United States.

References:

Ailshire, J., & Carr, D. (2021). Cross-National Comparisons of Social and Economic Contexts of Aging. *The Journals of Gerontology: Series B*, *76*(Supplement_1), S1–S4. https://doi.org/10.1093/geronb/gbab049

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to Meta-Analysis*. John Wiley & Sons.

Rai D., Zitko P., Jones K., Lynch J., & Araya R. (2013). Country- and individual-level socioeconomic determinants of depression: Multilevel cross-national comparison. *The British Journal of Psychiatry*, *202*(3), 195–203. https://doi.org/10.1192/bjp.bp.112.112482

COMMENT 1-15. *The authors need to theoretically justify their choice of cut-offs for effect modifiers (e.g., below vs above 65 for age).*

RESPONSE 1-15. Thank you very much for this suggestion. We provided the rationale for the selection of cut-off points, which was based on the official definition of the United Nations and prior literature.

The added part is copied below:

Methods Section, **Statistical analyses** Part: "The category of all modifiers except for age and chronic conditions was predefined (see Covariates). Age was categorized into 55–64 years and ≥65 years according to the definition of older people by the United Nations[90]. Chronic condition was classified as either no conditions or ≥1 condition, based on approximately one of the mean number of chronic conditions in our study (Supplementary Table 1) and in line with previous literature[4]."

COMMENT 1-16. *In Results, the authors should avoid interpretative and subjective language. For example, the authors used the term "effect" to refer to statistical associations. I believe the authors know the distinction between these two concepts but being precise on such a language will improve the manuscript. Similarly, the authors used the word "significantly" to refer to the effect sizes rather than statistical significance. This is a very subjective and vague word, and should be avoided.*

RESPONSE 1-16. Thank you so much for pointing out this issue. To avoid the subjective words, we first replaced the terms "effect" or "impact" with "association" or "relationship" in the revised manuscript, especially in the **Results** Section. Then, we rewrote all sentences that did not use "significantly" to refer to the statistical significance. The revised sentences are copied below:

Abstract: "These findings could provide important implications for public health policies and practices in promoting mental health in later life through the Internet, especially in countries with limited Internet access and mental health services."

Results Section, **Subpopulation analyses** Part: "After stratifying participants by the genetic risk (i.e., low [quintile 1], intermediate [quintile 2–4], and high [quintile 5]) of depression, Internet use was beneficially associated with depressive symptoms and self-reported health across three genetic risk categories in England and the USA."

REVIEWER 2'S COMMENTS

Remarks to the Author:

COMMENT 2-1. *This is a nice written manuscript investigated the association of internet use and depressive symptoms in people aged 50 and over from six aging cohorts. The association was consistent across the varied cohorts, and some sensitivity analyses also support the conclusion. Here are some comments for consideration.*

Did the authors analyze the data with reversing the x (depressive symptoms) and y (internet use)? Is this association could be bi-directional?

RESPONSE 2-1. Thank you very much for the comments. We agree with the reviewer on the need to characterize the potential bidirectional relationship between Internet use and depressive symptoms. Following the reviewer's suggestion, we fitted generalized estimating equation models with each of the three health indicators (i.e., depressive symptoms, life satisfaction, and self-reported health) as the independent variable and Internet use as the dependent variable to test the bidirectional association. Results showed that participants with fewer depressive symptoms, higher life satisfaction, and better self-reported health were more likely to use the Internet. This finding suggests that further studies are required to examine the causal ordering of relationships between Internet use and mental health. We copied the details of methods, results, and discussions below:

In the main text:

Methods Section, **Statistical analyses** Part: "Finally, we fitted generalized estimating equation models with each of the three health indicators as the independent variable and Internet use as the dependent variable to test the potential bidirectional association."

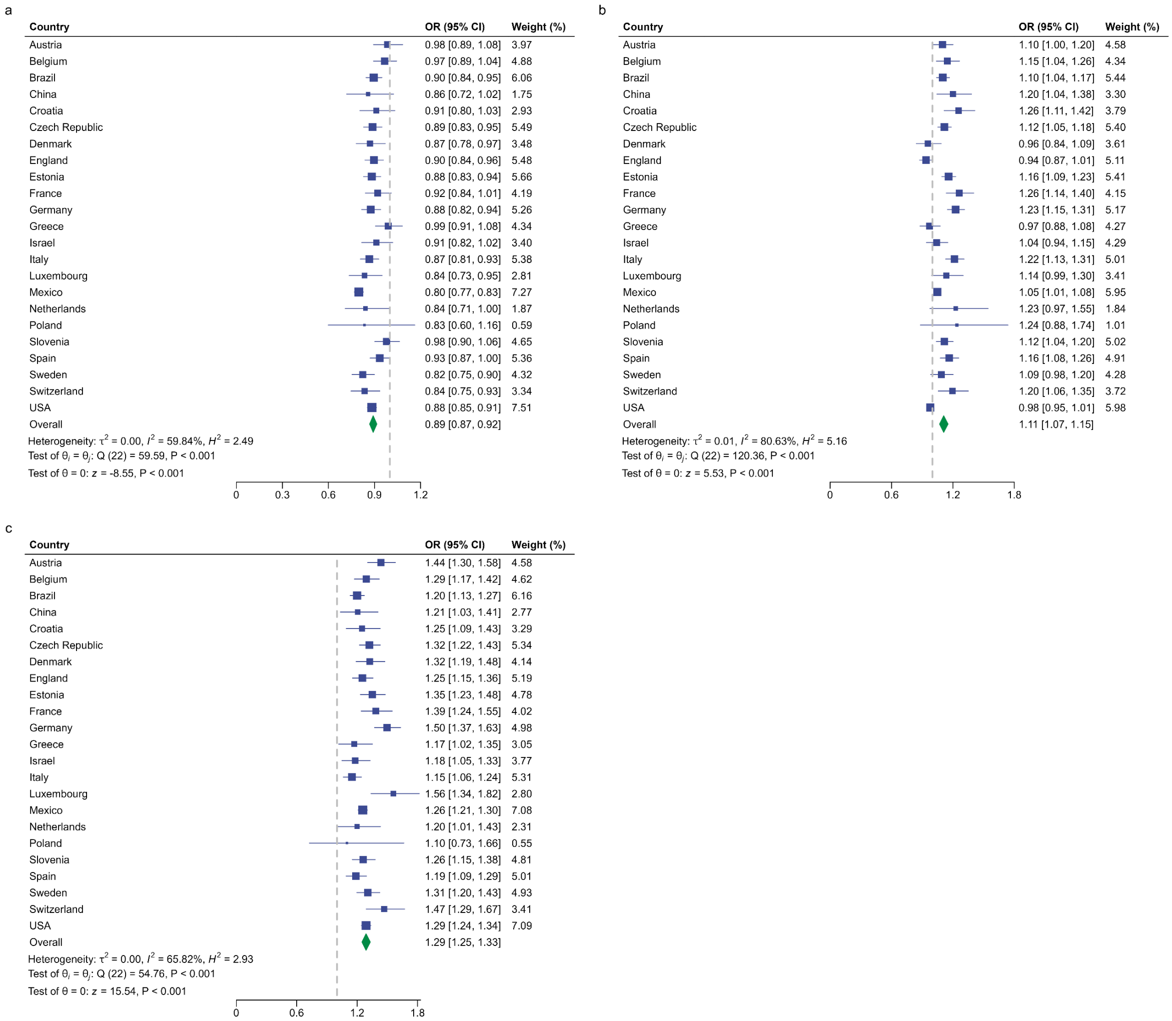
Results Section, **Sensitivity analyses** Part: "Finally, participants with fewer depressive symptoms, higher life satisfaction, and better self-reported health were more likely to use the Internet, suggesting a potential bidirectional association between Internet use and mental health (Supplementary Fig.20)."

Discussion Section, **Eleventh** Paragraph: "Finally, it should be noted that Internet usage is likely to have a bidirectional relationship with mental health, highlighting the need for future studies to examine this reciprocal association."

In the Supplementary Materials:

Supplementary Methods Section, **Bidirectional association between mental health and Internet use** Part: "To investigate the potential bidirectional association between mental health and Internet use, we performed estimating equation models (geepack package) with an exchangeable working correlation and each of three health indicators as the independent variable. Internet use was modeled as a binomial outcome with a logit link function. Among the participants aged ≥50 years at baseline, we excluded those who had missing information on Internet use and three outcomes at baseline, without the reassessment of Internet use at any follow-up surveys, had been diagnosed with memory-related diseases or psychological disorders at baseline, or had missing values on baseline covariates to generate analytic samples (n=90,417 from 23 countries)."

The added figure is provided below:



Supplementary Fig.20: Associations of depressive symptoms (a), life satisfaction (b), and self-reported health (c) with Internet use. Data are presented as the ORs with 95% CIs. The OR for each country (blue square) was separately estimated by the generalized estimating equation model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall OR (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. OR = odds ratio, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.

COMMENT 2-2. *Did the authors try to investigate the association of internet use and incident of depressive symptoms (none at baseline but occur in the follow-ups)?*

RESPONSE 2-2. Thank you so much for the valuable suggestion. Following the reviewer's suggestion, we first re-ran linear mixed models only including participants without depression at baseline. Then, we performed Cox proportional hazards models to examine the association between Internet use and the risk of incident depression. All analyses showed consistent results with our main findings, confirming the robustness of our results. Detailed methods, results, and discussions are copied below:

In the main text:

Methods Section, **Statistical analyses** Part: "Fourth, to mitigate the impact of pre-existing depressive symptoms, we re-ran the analyses by only including participants without depression at baseline. We also performed Cox proportional hazards models to examine the association between Internet use and the risk of incident depression."

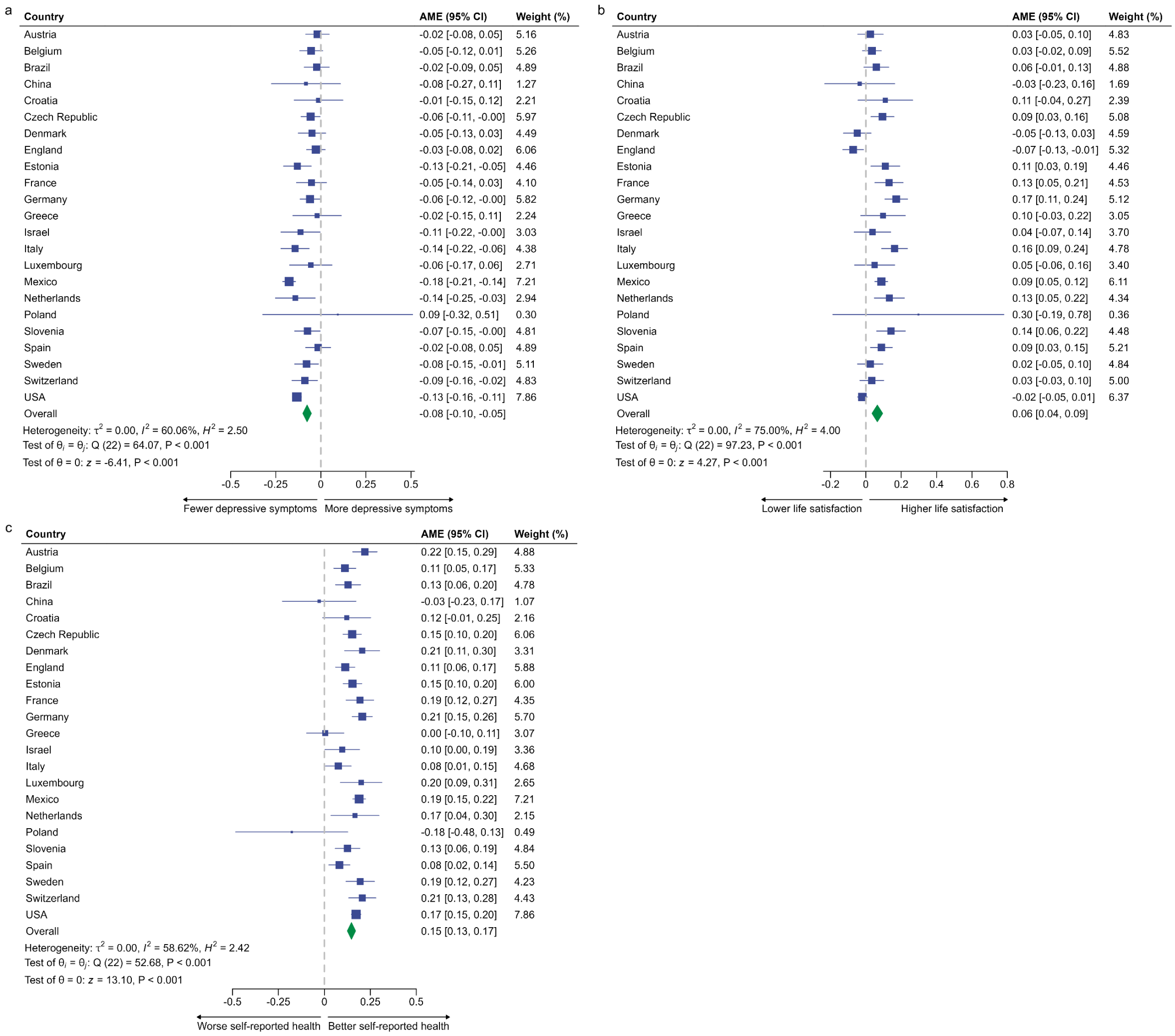
Results Section, **Sensitivity analyses** Part: "When restricting to participants without depression at baseline, the associations between Internet use and mental health showed little difference compared with the main analyses (Supplementary Fig.18). Furthermore, Internet use was associated with a reduced risk of incident depression (pooled hazard ratio = 0.85, 95% CI = 0.79 to 0.91, Supplementary Fig.19). As such, these data did not show evidence of reverse causality in our study."

Discussion Section, **Eleventh** Paragraph: "Our sensitivity analyses after considering the heterogeneity in Internet use measures, selection bias, and reverse causality separately showed that our findings were robust."

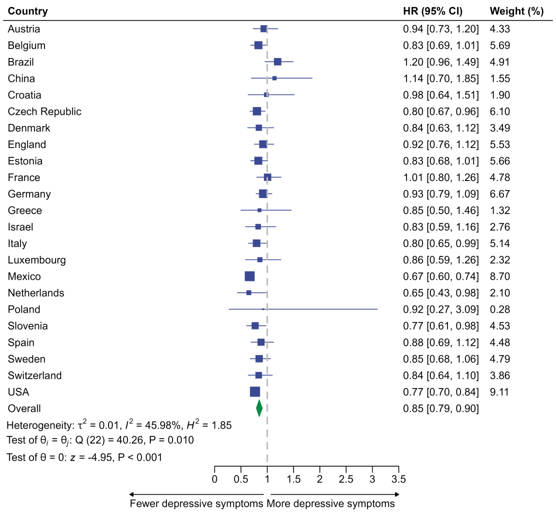
In the Supplementary Materials:

Supplementary Methods Section, **Association between Internet use and the risk of incident depression** Part: "To reduce the influence of existing depressive symptoms, we re-fitted covariate-adjusted linear mixed models among participants without depression at baseline (*n*=69,370 from 23 countries) to assess the association of Internet use with mental health. Depression was defined based on the recommended thresholds derived from validation work among older adults (Supplementary Table 10)[2–6]. Additionally, we performed Cox proportional hazards models (survival package) adjusted for all covariates to explore the relationship between Internet use and new onset depression. Participants were censored at the date of first occurrence of depression or the end of follow-up, whichever came first.”

The added figures are provided below:



Supplementary Fig.18: Associations of Internet use with depressive symptoms (a), life satisfaction (b), and self-reported health (c) among participants without depression at baseline. Data are presented as the AME with 95% CIs. The AME for each country (blue square) was separately estimated by the linear mixed model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall AME (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. AME = average marginal effect, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.



Supplementary Fig.19: Associations of Internet use with incident depression. Data are presented as the HRs with 95% CIs. The HR for each country (blue square) was separately estimated by the Cox proportional hazards model adjusted for age, gender, marital status, household size, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and IADL. The overall HR (green diamond) was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. HR = hazard ratio, CI = confidence interval, ADL = activities of daily living, IADL = instrumental activities of daily living.

COMMENT 2-3. *I have come across several highly cited paper which have already reported the very topic, what's the extra contribution, in addition to multi cohorts, for this one?*

Cotten S R, Ford G, Ford S, et al. Internet use and depression among older adults[J]. Computers in human behavior, 2012, 28(2): 496-499.

Cotten S R, Ford G, Ford S, et al. Internet use and depression among retired older adults in the United States: A longitudinal analysis[J]. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 2014, 69(5): 763-771.

Liao S, Zhou Y, Liu Y, et al. Variety, frequency, and type of Internet use and its association with risk of depression in middle-and older-aged Chinese: a cross-sectional study[J]. Journal of Affective Disorders, 2020, 273: 280-290.

RESPONSE 2-3. Thank you for providing these highly cited relevant papers, which are very helpful in motivating our study. We have included these papers in our references and added another recent review paper (Cotten et al., 2022). Compared to these studies, our study has two main contributions in addition to using multicohort data.

First, we quantified the portion of between-country heterogeneity explained by seven country-level factors. This analysis is important to inform public health policies on how the country can improve population-level mental health using the Internet, while it is often ignored in most cross-national research (Ailshire & Carr, 2021).

Then, we examined the relationship between Internet use and mental health in subgroups stratified by demographic characteristics, socioeconomic positions, health behaviors, physical health status, and genetic risks. In particular, by incorporating polygenic scores, we examined the interaction between Internet use and genetic risk for mental health. We also added the motivations for these analyses in the **Introduction** Section and the corresponding policy implications in the **Discussion** Section. However, compared to the study by Liao and colleagues (Liao et al., 2020), we only considered whether to use the Internet and the frequency and cumulative period of Internet use. Other information related to Internet usage, such as duration, purposes of use, types of online activities, and types of devices, needs further investigation. We have mentioned this limitation in the **Discussion** Section.

The added statements are shown below:
Introduction Section, **Third** Paragraph: "Additionally, genetic and environmental factors may interact with each other and affect the risk of mental disorders such as depression[31]. There is evidence that the association between media use/Internet use and mental health was attributed to genetic vulnerabilities in young adults[32,33]. However, few studies have explored the interaction between genetic risk and Internet use and its influence on mental health in older populations. Evaluating how genetic, sociodemographic, and behavioral factors influence the relationship between Internet use and mental health is crucial for pinpointing which subpopulations would most benefit from interventions involving Internet use[34]. Recognizing the varied impact of Internet use is important for precision mental health promotion, which currently faces challenges in systematically incorporating genetic, sociodemographic, and behavioral characteristics into the selection of mental health intervention strategies[34]."

Discussion Section, **Eighth** Paragraph: "Overall, the observed heterogeneity in our study indicates the need for precision mental health promotion that matches people to the interventions based on their genetic, sociodemographic, and behavioral factors. Take an example, using mobile health apps to encourage fitness and monitor wellness may be more beneficial for mental health in older adults without sufficient physical activity, while the interventions can be other formats, such as e-learning or online hobbies, for those who exercise regularly (exercise information can be obtained from these aging cohort surveys; i.e. relates to physical inactivity factor in our study)[57]. In the future, Internet-based mental health interventions could even be tailored to individual needs by developing different digital tools[58]. However, evaluating the effectiveness of Internet use in subpopulations of older adults may be challenging, and more robust evidence from randomized controlled trials is required."

Discussion Section, **Tenth** Paragraph: "However, these risk factors were not captured by our datasets. Future research should consider more details of Internet use, such as daily screen time, purposes of use, types of online activities, and the interactions with other users, to inform effective Internet use-related intervention strategies for mental health promotion[62]."

Discussion Section, **Twelfth** Paragraph: "Third, Internet use was ascertained by self-reported items, which might have led to recall and information bias. In addition, we measured Internet use based on a simple item. More detailed information on Internet use patterns can be considered in future studies."

References:

Ailshire, J., & Carr, D. (2021). Cross-National Comparisons of Social and Economic Contexts of Aging. *The Journals of Gerontology: Series B*, *76*(Supplement_1), S1–S4. https://doi.org/10.1093/geronb/gbab049

Cotten, S. R., Schuster, A. M., & Seifert, A. (2022). Social media use and well-being among older adults. *Current Opinion in Psychology*, *45*, 101293. https://doi.org/10.1016/j.copsyc.2021.12.005

Liao, S., Zhou, Y., Liu, Y., & Wang, R. (2020). Variety, frequency, and type of Internet use and its association with risk of depression in middle- and older-aged Chinese: A cross-sectional study. *Journal of Affective Disorders*, *273*, 280–290. https://doi.org/10.1016/j.jad.2020.04.022

COMMENT 2-4. *The authors have discussed the mechanisms between the internet use and depression, family connection, which is good, do we consider to control it in the regression* *and do subgroup analyses? In view of your explanation, the association could be more marked in those who contact family members or friends less frequently.*

RESPONSE 2-4. Thank you so much for the valuable suggestion. We agree with the reviewer that considering contact with family members or friends in our analyses is helpful for elucidating the potential mechanism between the Internet use and mental health. Following the reviewer's comment, we included *contact with others*, defined as whether the participants contact with any of their children, parents, relatives, or friends at least once a week, as a covariate in the linear mixed model and performed subgroup analyses. Detailed definitions, as a part of **Supplementary Table 10**, are copied below:

Variable	**Harmonized value**	**Measurements in six studies**					
		HRS	**ELSA**	**SHARE**	**MHAS**	**CHARLS**	**ELSI-Brazil**
Contact with others	Weekly contact	Whether the respondent contacts with any of their children, parents, relatives, and friends either in person or by phone, mail, or e-mail at least once a week in the past year? • Yes	Whether the respondent contacts with any of their children, relatives, and friends either in person or by phone, mail or e-mail, or by text at least once a week? • Yes	Whether the respondent contacts with any of their children and parents, either in person, by phone, mail, e-mail, sms or mms at least once a week? • Yes	Whether the respondent contacts with any of their children, relatives, and friends, either in person, by mail, or by telephone at least once a week? • Yes	Whether the respondent contacts with any of their parents in person, or children either in person, by mail or e-mail at least once a week? • Yes	In the past 12 months, have you kept contact with other people by letters, telephone, email, and/or social networks on the Internet? • Yes
	Less than weekly contact	• No	• No	• No	• No	• No	• No

The results after adjusting for *contact with others* were largely consistent with the original results. Moreover, consistent with our hypothesis, we observed a stronger protective association of Internet use with depressive symptoms and life satisfaction among individuals with less than weekly contact than those with weekly contact (overall average marginal effect: -0.12 vs -0.09 in depressive symptoms; 0.06 vs 0.11 in life satisfaction). However, these differences were not statistically significant. The tables that summarize the effect sizes in subpopulation analyses are too long to be presented in the response letter. Please see

Supplementary Tables 6–8 in the Supplementary Materials for details. We added a statement in the **Discussion** Section to highlight our findings as supporting evidence for the potential mechanism between Internet use and mental health — Internet use can facilitate connections with family and friends and expand their social networks among middle-aged and older adults, then reducing their loneliness and social isolation, and thereby contributing to better mental health.

The modified figure for the subpopulation analyses is copied below:

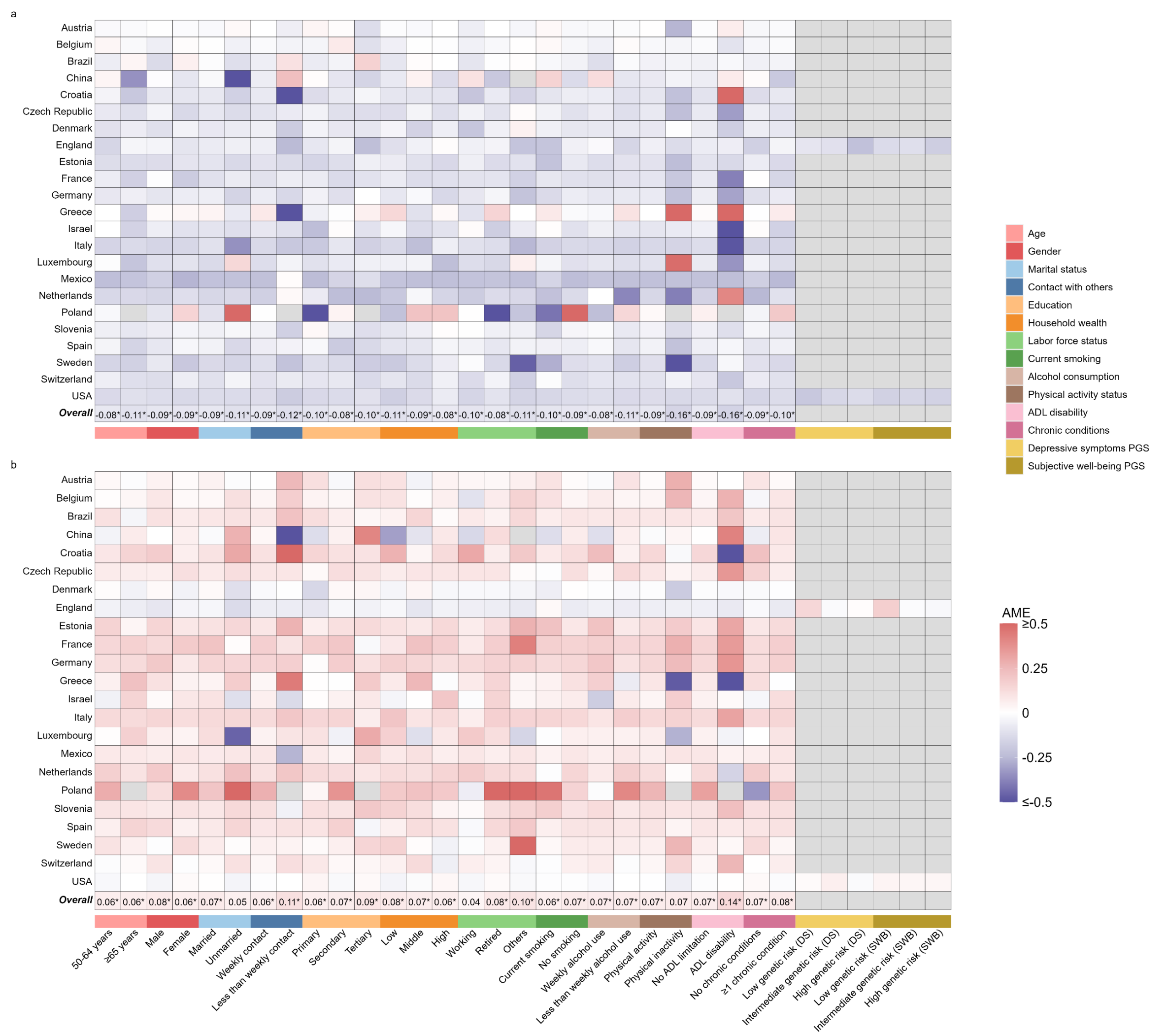


Fig.2: Associations of Internet use with depressive symptoms (a) and life satisfaction (b) by age, gender, marital status, contact with others, education, household wealth, labor force status, current smoking, alcohol consumption, physical activity, chronic conditions, ADL, and genetic risks assessed by PGS for depressive symptoms and subjective well-being. The AME for each country was separately estimated by the linear mixed model adjusted for all covariates (not including genetic risks) except the stratification variable. The overall AME was pooled using the random-effect meta-analysis with a restricted maximum likelihood estimator. The color of cells indicates the effect sizes (AME) between internet use and mental

health outcomes. Grey rectangles indicate missing effect sizes because the model failed to converge with the limited sample size of subpopulations in China and Poland, as well as the genetic data is not available in all countries except for England and USA. AME = average marginal effect, ADL = activities of daily living, PGS = polygenic scores, DS = depressive symptoms, SWB = subjective well-being.

The added statement is copied below:

Discussion Section, **Second** Paragraph: “This explanation was partly supported by our subgroup analyses showing a more pronounced beneficial association between Internet use and mental health in older people who have less frequent contact with others.”

REVIEWER 3’S COMMENTS

Remarks to the Author:

COMMENT 3-1. *The paper provides a good narrative on the association between Internet use and age-related mental health challenges in older adults. The data from six aging cohorts across 23 high- and middle-income countries is well-defined.*

As the authors have described, the protective role of Internet use varies in different subpopulations within a country, an important factor to consider when reviewing the social determinants of health. However, we need to consider the question of whether a higher frequency of Internet use is related to improved mental health. These findings could provide significant implications for public health policies and practices in promoting mental well-being in later life through Internet usage. This paper is well-written and contains useful information. I have a few minor points the authors may wish to consider.

The authors should consider reporting why they chose these datasets only.

RESPONSE 3-1. We thank the reviewer’s supportive comments. We added the reason for the selection of these cohorts in the **Methods** Section, **Study population** Part: “We selected these cohorts from the list provided by the Gateway to Global Aging Data platform (https://g2aging.org/survey/overview) based on two criteria: 1) availability of Internet use information and 2) having at least two waves of follow-ups.”

COMMENT 3-2. *Include internet usage as social support for older adults in measurable terms.*

RESPONSE 3-2. Thank you for the suggestion. Following the reviewer's comment, we quantified its impact by the improvement in three health indicators. We added a sentence in the **Discussion** Section, **Second** Paragraph to present the results more clearly: "Specifically, we found that using the Internet was associated with 0.09 decreased depressive symptom scores, 0.07 incremental life satisfaction scores, and 0.15 increased self-reported health scores in older adults."

COMMENT 3-3. *Explain the negative association (i.e., more beneficial) between Internet use and depressive symptoms in Fig.3a.*

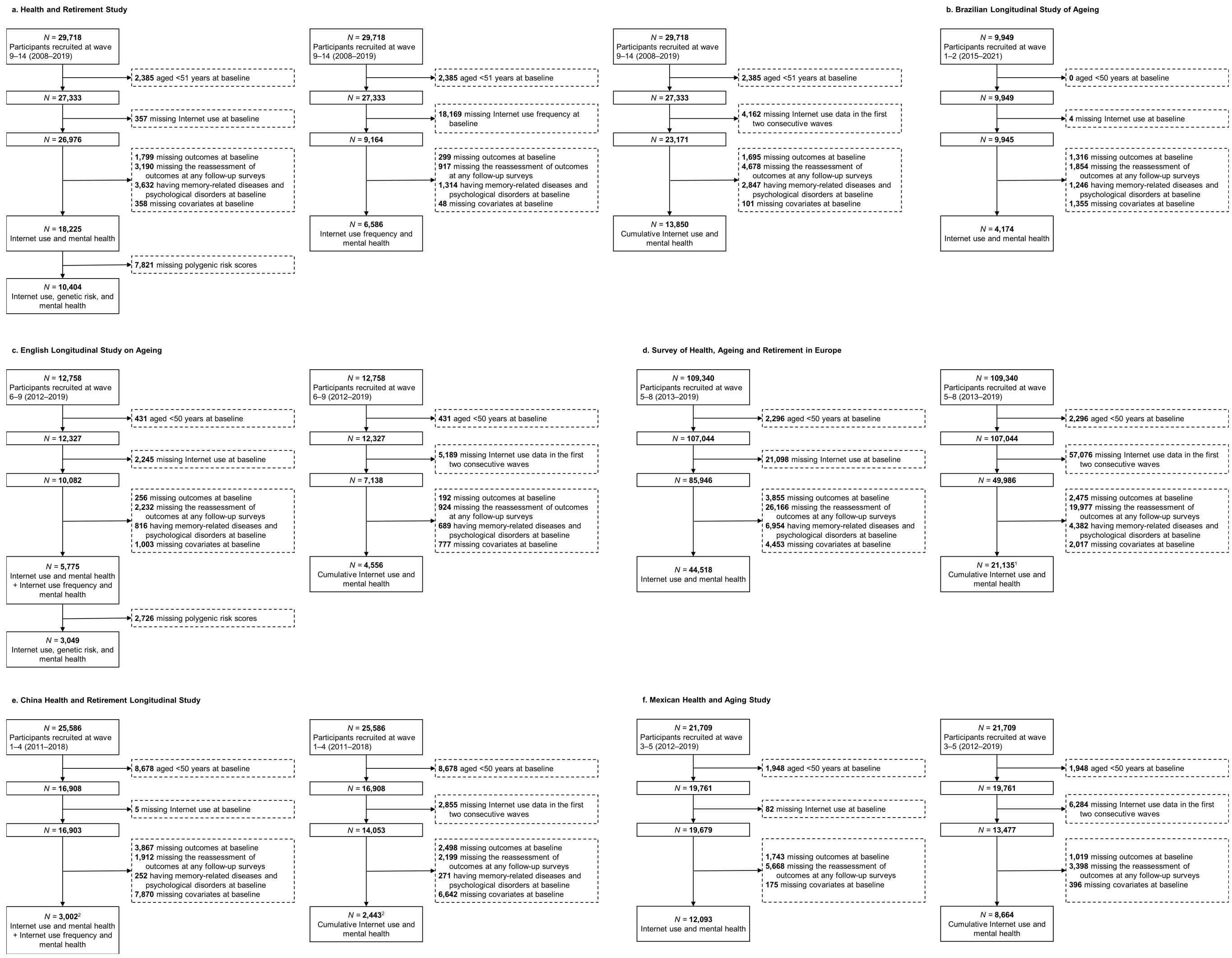
RESPONSE 3-3. We thank the reviewer for the important comment. Following the suggestion, in the revised manuscript, we expanded the discussion on the more beneficial association between Internet use and depressive symptoms in subpopulations. The modified paragraph is copied below:

Discussion Section, **Seventh and Eighth** Paragraphs: "In addition to cross-national variations, there were differences in the associations of Internet use with mental health between subpopulations within a country. Specifically, our results showed that individuals with advanced age, physical inactivity, and ADL disability may benefit more from using the Internet, which is in line with prior findings[26–28]. A plausible explanation is that these vulnerable people are at higher risk of developing mental health problems, and Internet use may serve a more efficient compensatory function[12]. We also found a stronger protective role of Internet use among adults with high levels of education. The results may be explained by the high digital and health literacy of these populations, which could help them access and use online healthcare information and services for health management and avoid problematic Internet use and online scams[51,55]. These findings, together with another recent empirical study[56], implicate that facilitating digital inclusion for this population requires increasing access to the Internet and related devices, as well as enhancing their ability, namely digital and health literacies, to effectively use the Internet for promoting their own health and well-being[48].

Moreover, our study showed that Internet use could improve mental health among middle-aged and older adults in England and the USA across each level of genetic risk, suggesting that using the Internet can partially reduce the genetically determined increased risk of mental health problems. Overall, the observed heterogeneity in our study indicates the need for precision mental health promotion that matches people to the interventions based on their genetic, sociodemographic, and behavioral factors. Take an example, using mobile health apps to encourage fitness and monitor wellness may be more beneficial for mental health in older adults without sufficient physical activity, while the interventions can be other formats, such as e-learning or online hobbies, for those who exercise regularly (exercise information can be obtained from these aging cohort surveys; i.e. relates to physical inactivity factor in our study)[57]. In the future, Internet-based mental health interventions could even be tailored to individual needs by developing different digital tools[58]. However, evaluating the effectiveness of Internet use in subpopulations of older adults may be challenging, and more robust evidence from randomized controlled trials is required."

COMMENT 3-4. *Sample selection procedures for each cohort need further clarification in Fig.1.*

RESPONSE 3-4. Thank you so much for this suggestion. We modified the sample selection procedures by adding the number of excluded participants for each exclusion criterion. The updated figure is provided in **Extended Data Fig.1** in the main text and copied below:



Extended Data Fig.1: Flowchart of sample selection in the Health and Retirement Study [HRS](a),-Brazilian Longitudinal Study of Aging [ELSI](b), English Longitudinal Study on Ageing [ELSA](c), Survey of Health, Ageing and Retirement in Europe [SHARE](d), China Health and Retirement Longitudinal Study [CHARLS](e), and Mexican Health and Aging Study [MHAS](f). In the present study, outcomes included depressive symptoms, self-reported health, and life satisfaction. Memory-related diseases were Alzheimer's disease, dementia, Parkinson's disease, and other memory-related conditions. Psychological disorders were emotional, nervous, and psychiatric problems. Memory-related diseases and psychological disorders were not available in MHAS. [1]Greece and Poland were excluded from examining associations between cumulative Internet use and well-being due to the limited sample size of populations (only two participants from each country). [2]A large proportion of participants were excluded due to the missingness of household wealth and physical activity. In CHARLS, a random half sample was selected to ask questions on physical activity.

COMMENT 3-5. *In the introduction section, one would expect a coherent exposition on well-being defined as high and low, moderate, or medium.*

RESPONSE 3-5. Thank you very much for this valuable comment. We acknowledge the need to provide more theoretical backgrounds of mental health and well-being. Therefore, in the **Introduction** Section, we introduced the definition of mental health proposed by the World Health Organization (WHO) and the staging model for mental health developed by The Lancet Commission on Global Mental Health and Sustainable Development (World Health Organization, 2022; Patel et al., 2018). WHO defines mental health as a state of well-being that enables individuals to cope with the stresses of life, realize their abilities, learn and work well, and contribute to their community, suggesting that mental health is more than the absence of mental disorders. According to this staging model, mental health is a complex continuum covering from well-being (e.g., life satisfaction) to conditions (e.g., depressive symptoms). These concepts highlight that preventing mental disorders and enhancing mental well-being are both essential in mental health promotion. According to the staging model, "mental health" refers to the full spectrum of mental health states, "mental illness" indicates pathological disease states, and "mental well-being" acts as the positive component of mental health. So, we replaced the term "mental well-being" with "mental health" in the revised manuscript.

We copied the added parts in the main text below:

Introduction Section, **First** Paragraph: "Mental health, as defined by the World Health Organization (WHO), is a state of well-being that enables individuals to manage the stresses of life, realize their abilities, learn and work well, and contribute to their community[2]. This definition extends beyond the absence of mental disorders. According to the staging model for mental health developed by The Lancet Commission on Global Mental Health and Sustainable Development, mental health encompasses a broad spectrum, ranging from well-being (e.g., life satisfaction) to various conditions (e.g., depressive symptoms)[2,3]. These concepts advocate for mental health promotion, emphasizing the need to not only prevent mental disorders but also enhance mental well-being."

References:

Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., Chisholm, D., Collins, P. Y., Cooper, J. L., Eaton, J., Herrman, H., Herzallah, M. M., Huang, Y., Jordans, M. J. D., Kleinman, A., Medina-Mora, M. E., Morgan, E., Niaz, U., Omigbodun, O., … UnÜtzer, Jü. (2018). The Lancet Commission on global mental health and sustainable development. *Lancet*, *392*(10157), 1553–1598. https://doi.org/10.1016/S0140-6736(18)31612-X

World Health Organization. (2022). *Mental health*. https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response

Dear Editor and Reviewers,

Thank you so much for taking the time to review this manuscript. We really appreciate all your comments and suggestions! We hope that the revised manuscript could meet the requirements for Nature Human Behaviour. All revised portions are marked in red in the revised manuscript.

The main comments and our specific responses are detailed below.

REVIEWER 1'S COMMENTS

Remarks to the Author:

COMMENT 1-1. *The addition of the equation for the main analysis is helpful. It seems that the authors are essentially looking at cross-sectional associations between internet use and the outcomes. In their "longitudinal" model, internet use was assessed at the same time point as the outcome assessment. The model allowed this association to vary over time (via the interaction term with time). The time-updated cross-sectional associations were then pooled as AME. Hence, much of the observed associations are likely due to reverse causation due to the lack of temporal ordering between the internet use and the outcomes. The authors did a sensitivity analysis excluding those with baseline depression, but this will address reverse causation for the "baseline" internet use only. Subsequent cross-sectional associations remain subject to reverse causation. Nonetheless, the paper overall emphasized the longitudinal nature of the analysis.*

RESPONSE 1-1. Thank you very much for the comment. Follow the reviewer's suggestion, we removed the term "longitudinal" in the revised manuscript. The revised sentences are copied below:

Abstract: "We used linear mixed models and meta-analyses to examine the association between Internet use and mental health among 87,559 adults aged ≥50 years from 23 countries."

Introduction Section, **Second** Paragraph: "This highlights the need for cross-national, population-based studies to examine if and how Internet use can benefit mental health in middle-aged and older adults, despite negative results being observed among younger persons instead[30].”

Introduction Section, **Fourth** Paragraph: “Therefore, by integrating harmonized data from six aging cohorts across 23 countries and two-stage individual participant data (IPD) meta-analysis …”

Results Section, **Study population** Part: “We developed linear mixed models and performed meta-analyses using six aging cohorts from 23 high- and middle-income countries …”

Results Section, **Baseline Internet use and mental health** Part: “Associations between baseline Internet use and subsequent mental health are shown in Fig.1 and Supplementary Fig.1.”

Discussion Section, **First** Paragraph: “In this cross-national study, we found that baseline Internet use was associated with subsequent reduced depressive symptoms, higher life satisfaction, and better self-reported health among adults aged ≥50 years across 23 countries.”

Discussion Section, **Tenth** Paragraph: “The main strength of our study is the use of harmonized data from six aging cohorts across 23 countries.”

Discussion Section, **Eleventh** Paragraph: “First, the possibility of residual confounding and reverse causation means we cannot establish causality between Internet use and mental health.”

Methods Section, **Study population** Part: “Six aging cohorts included in this study were HRS, ELSA, SHARE, CHARLS, MHAS, and ELSI.”

Methods Section, **Ethics and inclusion statement** Part: “This study analyzed six aging cohorts across the USA, England, Europe, China, Mexico, and Brazil.”

COMMENT 1-2. *I still believe the country-level factors analysis is uninformative and unnecessary. It is purely descriptive, yet the authors interpreted the meta-regression results* *as if the between-country variation can be attributed to each specific country-level indicator. They used language such as "attributed" or "explain". Identifying the source of the between-country heterogeneity would be a whole another paper, and I suggest the authors keep the paper more focused on identifying the heterogeneity itself. The paper is already full of information, and adding extra information that is very exploratory obscures the key contribution of the study (and it is not reader-friendly).*

RESPONSE 1-2. We thank the reviewer for the comments. Following the reviewer's suggestion, we first removed the results of the country-level factors analysis from our main findings in the **Abstract** and **Discussion** section. Then, we only kept brief descriptions in the **Methods** and **Results** sections and one short paragraph in the **Discussion** section. Detailed methods are provided in the **Supplementary Methods**. Finally, we replaced the term "attributed" or "explain" in the country-level factors analysis with "correlated with". The revised parts are copied below:

Results Section, B**aseline Internet use and mental health** Part: "Our results showed that the association between Internet use and mental health outcomes was correlated with country-level factors, such as the Gini index and digital skills among active population (Supplementary Fig.6). The correlation between the prevalence of Internet use in our samples and seven country-level factors is given in Supplementary Fig.5."

Discussion Section, **Fifth** Paragraph: "Nevertheless, we also observed some variations in the relationship between Internet use and mental health across countries. Our meta-regression analyses revealed nonlinear associations of the effect sizes of Internet use for depressive symptoms and life satisfaction with the Gini index, healthy life expectancy (HALE) at birth, Internet use rate, and digital skill levels. Additionally, we found that the positive correlation between Internet use and self-reported health increased with the GDP per capita and world happiness index. Although these findings are exploratory, they could help generate hypotheses for future studies to investigate the cross-national differences observed in our study."

Methods Section, **Statistical analyses** Part: "We further conducted meta-regressions to explore the heterogeneity variance in the meta-analysis with seven country-level factors, including GDP per capita, Gini index, world happiness index, Hofstede's index, Internet use rate, digital skills among active population, and HALE at birth, as the independent variable separately (Supplementary Methods)."

COMMENT 1-3. *The authors are using an incorrect approach for pooling estimates after multiple imputation. Specifically, confidence intervals need to be pooled via Rubin's rule and are not simple mean of lower/upper bounds from each imputed dataset.*

RESPONSE 1-3. Thank you so much for pointing out this issue. In the previous version of the manuscript, we applied the Rubin's rules to pool the estimates for all analyses except for the parametric g-formula. Pooling results for the parametric g-formula is more complicated because it requires bootstrapping to calculate standard errors. In the revised manuscript, we used the "MI Boot" method, recommended by Schomaker et al (Schomaker & Heumann, 2018), to obtain final estimates with confidence intervals. Specifically, we generated 10 imputed datasets. Within each imputed dataset, we calculated the point estimate using imputed, but not yet bootstrapped data and bootstrapped 100 times to derive the standard error. Finally, we pooled 10 point estimates and standard errors, and generated standard multiple imputation confidence intervals using Rubin's rules. The updated results were consistent with our main findings. Additionally, we provided the formula used for pooling estimates after multiple imputation in the Methods Section of the updated manuscript. The revised parts are copied below:

Methods Section, **Statistical analyses** Part: "Following Rubin's rules, the pooled point estimate can be calculated as equation (1):

$$\bar{Q} = \frac{1}{m}\sum_{i=1}^{m} \hat{Q}_i \tag{1}$$

where $\hat{Q}_i$ refers to the point estimate from the imputed dataset i, $i = 1, \dots, m$. In our study, $m = 10$. Total variance can be obtained using the between imputation variance $B = \frac{1}{m-1}\sum_{i=1}^{m}\left(\hat{Q}_i - \bar{Q}\right)^2$ and the average within imputation variance $\bar{U} = \frac{1}{m}\sum_{i=1}^{m}\widehat{Var}(\hat{Q}_i)$ with equation (2):

$$T = \bar{U} + \left(1 + \frac{1}{m}\right)B \tag{2}$$

Finally, we calculated the confidence interval for pool estimates using equation (3):

$$\bar{Q} \pm t_v\left(\frac{\alpha}{2}\right)\sqrt{T} \quad (3)$$

where t_v is the $\alpha/2$ quantile of the t distribution with v degrees of freedom."

Supplementary Methods Section, **Effect of cumulative Internet use on mental health** Part: "Within each imputed dataset, we calculated the point estimate using imputed, but not yet bootstrapped data and bootstrapped 100 times to derive the standard error. Finally, we pooled 10 point estimates and standard errors, and generated standard multiple imputation confidence intervals using Rubin's rules. This algorithm has been recommended especially when standard errors cannot be calculated easily for causal inference estimators such as the g-formula. Details of the algorithm can be found in the Schomaker et al.'s work (see "Method 2, MI Boot" part)."

The updated figures are copied below:

In the main text:

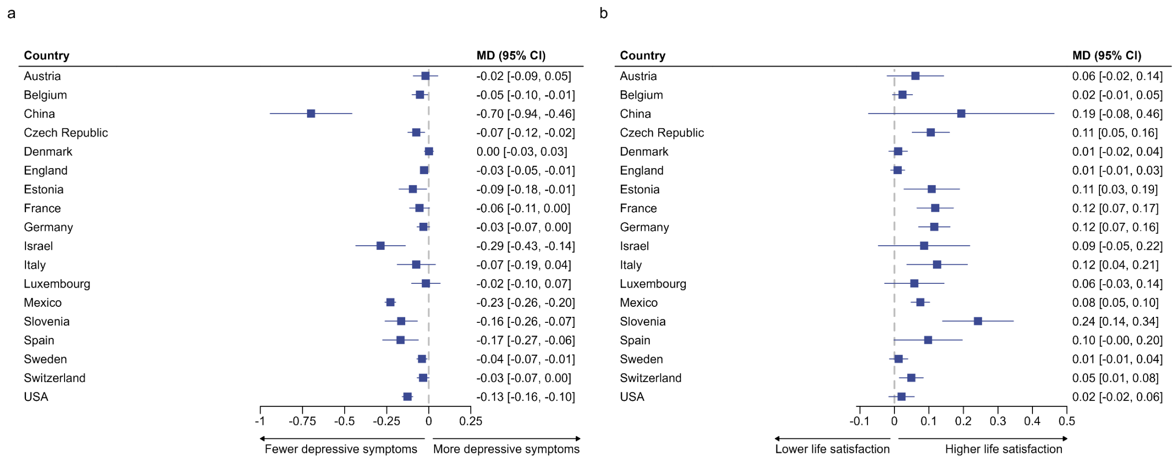
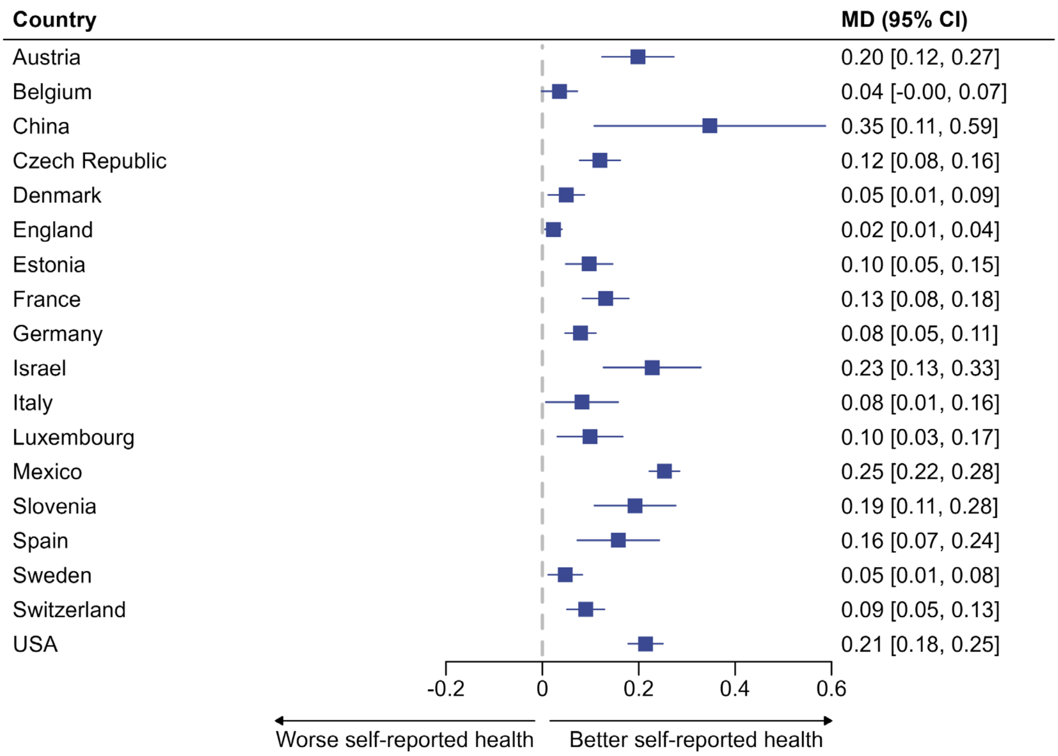


Fig.5: Estimated mean difference in the standardized scores of depressive symptoms (a) and life satisfaction (b) in the last follow-up between populations following the natural course (as the reference group) and those who received the intervention consistently (*n*=50,644 from 18 countries). Data are presented as the MDs with 95% CIs. MD = mean difference, CI = confidence intervals.



Supplementary Fig.11: Estimated mean difference in the standardized scores of self-reported health in the last follow-up between populations following the natural course (as the reference group) and those who received the intervention consistently (*n*=50,644 from 18 countries). Data are presented as the MDs with 95% CIs. MD = mean difference, CI = confidence intervals.

Reference:

Schomaker, M., & Heumann, C. (2018). Bootstrap Inference When Using Multiple Imputation. *Statistics in Medicine*, *37*(14), 2252–2266. https://doi.org/10.1002/sim.7654