medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Long-Term Impact of Stressful Life Events on Breast Cancer Risk: A 36-Year Genetically Informed Prospective Study in the Finnish Twin Cohort

Elissar Azzi[1*], Hannes Bode[1,2*], Teemu Palviainen[1], Mikaela Hukkanen[1,2], Miina Ollikainen[1,2], Jaakko Kaprio[1]**

1 Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland
2 Minerva Foundation Institute for Medical Research, Helsinki, Finland

*Shared first authorship. These authors contributed equally to this work.

**Correspondence to Prof. Jaakko Kaprio (e-mail: jaakko.kaprio@helsinki.fi), Institute for Molecular Medicine Finland, Tukholmankatu 8, 00290 Helsinki, Finland

ORCID:

Elissar Azzi: 0009-0004-1091-7346
Hannes Frederik Bode: 0000-0003-3232-6518
Teemu Palviainen: 0000-0002-7847-8384
Mikaela Hukkanen: 0000-0002-8574-8274
Miina Ollikainen: 0000-0003-3661-7400
Jaakko Kaprio: 0000-0002-3716-2455

Keywords:

Stressful life events, breast cancer, cohort study, twins, polygenic risk score, methylation

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

ABSTRACT

To investigate potential interactions between stress, genetic predisposition, and breast cancer, we employed a multifaceted approach, analysing polygenic risk scores (PRS), DNA methylation (DNAm) patterns, stressful life events, and breast cancer risk using data from 10342 women in the Finnish Twin Cohort, including 719 breast cancer cases. Stressful life events data was gathered from a 1981 questionnaire while cancer and mortality data were obtained from Finnish Cancer Registry, Finnish Population Register and Statistics Finland. DNAm data from breast cancer discordant twin pairs was used to explore within-pair differences between 212 breast cancer associated CpG sites and stressful life events. A hazard ratio (HR) of 1.05 (95%CI 1.02-1.08) was observed for breast cancer per one-event increase in total stressful life events. Women exposed to over 10 stressful life events exhibited an almost two-fold increase in risk of breast cancer (HR=1.93, 95%CI 1.19-3.14). Stratifying by birth decade revealed a lower HR for those born after 1950 (HR 1.01; 95%CI 0.96-1.06) compared to those born between 1940-1949 (HR 1.07; 95%CI 1.01-1.12). In a subsample with genetic data, no association with breast cancer PRS was seen (HR = 1.03, 95%CI 0.92-1.15 per SD of PRS). Stressful life events exposure associated with breast cancer related DNAm in 42 of the 212 CpG sites. This 36-year genetically informed study highlights the long-term association between stressful life events and breast cancer risk. These findings suggest epigenetic changes, but not polygenic risk may mediate the impact the effects of stressful life events on breast cancer incidence.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

INTRODUCTION

Breast cancer remains the predominant and lethal form of malignancy affecting women worldwide, with 2.3 million new cases diagnosed globally in 2020 alone [1]. Twin studies reveal that 31% of variability in risk is attributed to genetic factors [2], encompassing well-known high-impact genes like *BRCA1/2*, *CHECK2*, and *PALB2* [3], as well as numerous genes with single nucleotide polymorphisms (SNPs), collectively quantifiable in a polygenic risk score (PRS) [4]. While significant progress has been made in understanding the genetic underpinnings of breast cancer, the full picture of its aetiology remains a work in progress. Environmental factors and health-related behaviours, including established ones such as alcohol consumption, obesity, female reproductive factors, and hormone exposures, contribute significantly to breast cancer risk [5].

Stress – the physical and emotional response to challenging or demanding situations– has potential implications for later life health through various physiological mechanisms. Chronic or repeated stress disrupts the hypothalamic-pituitary-adrenal (HPA) axis which can lead to elevated cortisol levels, a hormone essential in mammary gland function and implicated in promoting cell proliferation and tumour growth [6]. Stress weakens the immune system, hindering its ability to eliminate abnormal cells via immune evasion or changes in DNA repair mechanisms [7], and is possibly reflected in epigenetic profiles [8].

The link between stressful life events and breast cancer risk is a complex one, with ongoing epidemiological research exploring the potential connection yielding diverse results. While some studies suggest a positive association between stressful life events and breast cancer risk, others, such as the UK study of 106000 women [9] report no consistent evidence for an association of breast cancer risk with perceived stress levels or adverse life events in the preceding 5 years. A 2019 meta-analysis, drawing data from prospective and retrospective cohort studies assessing the impact of stress, or lack thereof, on breast cancer incidence published before 2018, provided a broader perspective highlighting a moderate overall association between stressful life events and breast cancer risk [(pooled risk ratio: 1.11 (95%CI 1.03-1.19)] [10]. A previous publication by Lillberg et al. (2003) on the Finnish Twin Cohort demonstrated an association between stressful life events and a heightened risk of incident breast cancer [11].

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

There is also a familial and genetic basis to many stressful life events, reflecting the biological basis of reactions to external stressors [12, 13]. Thus, a possible explanation for the association between stressful life events and breast cancer risk may be shared genetic predisposition. This hypothesis can be evaluated by examining monozygotic (MZ) twin pairs, in which the twin sisters differ both in their experience of stressful life events and breast cancer occurrence. Alternatively, a PRS can be constructed, based on numerous low-impact genetic variants across the genome identified from genome-wide association studies, to estimate inherent susceptibility to breast cancer. By examining the associations of breast cancer PRS with stressful life events, presence or absence of shared genetic predisposition can be tested.

The association between stressful life events years or decades prior to breast cancer diagnosis may be reflected in changes in DNA methylation. DNA methylation is a dynamic epigenetic modification, acting thus as a potential mediator between genotypes, environmental stressors and gene expression. Exposures such as smoking and experiences such as stressful life events encountered throughout the lifespan, can induce significant epigenetic variation through mechanisms such as altered DNA methylation patterns and chromatin remodelling [14]. This emerging understanding of epigenetic plasticity holds relevance in the context of stress and breast cancer. Adverse experiences have been implicated in modulating these epigenetic pathways within breast tissue, potentially leading to altered gene expression profiles that favour tumorigenesis [15, 16]. In Bode et al. (2024) [17] it was shown that 212DNA methylation sites are associated specifically with overall environmental risk for breast cancer on average 11 years prior to diagnosis among Finnish twin pairs discordant for breast cancer.

Our current investigation extends the study by Lillberg et al (2003) by 22 years, spanning 36 years to examine the association of the number of stressful life events with breast cancer risk. The analysis includes a comprehensive assessment of established and proposed risk factors as well as analysis of genetic and epigenetic information. We conducted standard cohort analyses followed by within-family analyses in MZ and dizygotic (DZ) twin pairs concordant and discordant for breast cancer occurrence. Leveraging the extensive range of birth years within our cohort, the analysis was further enriched by stratification by birth decade. To deepen our understanding of the association between stressful life events and breast cancer risk, we

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

employed a multifaceted approach, whereby PRS and DNA methylation patterns were examined to evaluate the relationships between stressful life events, genetic predisposition and epigenetic methylation marks and breast cancer incidence.

Identifying DNA methylation sites associated with both environmental breast cancer risk and exposure to stressful life events in our cohort could provide crucial insights into how stress experiences might influence breast cancer risk through long-term epigenetic modifications. By analysing PRS and DNA methylation patterns alongside exposure to stressful life events, this study aims to provide a more holistic view on breast cancer risk.

METHODS

Stressful Life Events Analysis

Twin cohort description

The data used in this study originates from the older Finnish twin cohort, consisting of same-sex twin pairs born before 1958 [18]. This cohort was followed up with comprehensive health and lifestyle questionnaires in 1975, 1981, 1990 and 2011. Of the 16160 female twins in the cohort, 12986 returned a questionnaire in the 1981 survey. The questionnaire was mailed in the autumn of 1981, with reminders into early 1982. The first questionnaires were returned on October 7, 1981, with a median return date of November 4, 1981. Follow-up for each person started from the date of return of the questionnaire and ended on December 31, 2018. After excluding those individuals with prevalent breast cancer or residence outside Finland, missing life event information and other covariates, final sample size was 10342, with 719 incident breast cancers from 1982 to 2018.

Risk factors

Primary risk factor data was obtained from the 1981 questionnaire, supplemented by the 1975 questionnaire where data was missing. The main risk factor of interest was the number of stressful life events (SLE score) in the past five years. The compilation of data on life events is explained in more detail by Lillberg et al. (2003) [11]. Essentially, the 1981 questionnaire contained questions on 17 major life events, including change of residence, loss of job, death of a relative and others [7]. These were aggregated for analysis on an individual basis, with the

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

number of events exceeding 10 reduced to 10, as persons with more than 10 events were few in number.

Other risk factors of interest were zygosity, marital status, oral contraceptive use, social class, BMI (calculated as the mean between 1975 and 1981), alcohol consumption (expressed in ethanol gram per day and categorised as abstainer, 1-10, 11-20 and 21 or more grams per day), smoking (never, former, current), and leisure time physical activity (quintiles, MET-hours per week). In addition, we assessed the link between breast cancer and reproductive history using latent classes of lifelong pregnancy data obtained from the 1975 questionnaire and the Finnish Population Register (hereafter "reproductive class" described in Hukkanen et al., 2024 [19]).

Outcome

Data on cancer incidence up to the end of 2018 was obtained from the Finnish Cancer Registry (FinData permit THL/6353/14.06.00/2023), while data on mortality and emigration were obtained from the Finnish Population Register and Statistics Finland [18].

Cases were defined as women who did not have breast cancer prior to response to the 1981 questionnaire but were diagnosed with breast cancer thereafter. In contrast, controls were defined as individuals who were not diagnosed with breast cancer until the time of censoring, due to emigration, death or end of the study in 2018. The cumulative hazard by age is shown in Supplemental Figure 1, with an overall cumulative incidence of 7%.

Evaluating the association of life events with breast cancer risk using polygenic risk scores

Genotypes for evaluating genetic breast cancer risk through PRS was available for a sub-sample of 4,601 women (4,237 controls and 364 cases, disease prevalence 7.9%). Within this sample,176 DZ twin pairs were discordant for breast cancer. A breast cancer PRS was derived from breast cancer GWAS summary statistics [20] for 1,069,529 Single Nucleotide Polymorphisms (SNPs) [4]. For details see Supplemental Material: Genetic Analyses.

Evaluating the association of life events with breast cancer risk using DNA methylation

Altogether 319 twin pairs (172 MZ and 147 DZ pairs) had information on DNA methylation data for 212 CpG sites related to breast cancer risk and the 1981 questionnaire SLE score. DNAm, was generated by Illumina Infinium MethylationEPIC (EPIC) platform (Illumina, San Diego,

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

CA, USA). Quality control and preprocessing of the data followed our in-house pipeline [17]. Bode et al. (2024) [17] identified 212 CpG sites in genome-wide blood DNAm data associated with future breast cancer risk, based on a discordant twin design of twin pairs in which one twin sister developed breast cancer later and the other twin sister was unaffected. This approach controls for technical confounding factors, age and shared environment, as well as germline genetics [21]. These results provide support for the view that the identified CpG sites' association with breast cancer risk in survival analysis likely reflects within pair differences in environmental exposures that increase breast cancer risk [17]. For details see Supplemental Material: DNA Methylation Analyses. To avoid confounding by breast cancer or related factors, twin pairs where either of the twins had a breast cancer diagnosis before DNA sampling were excluded. No exclusions were made based on other diseases occurring before or after DNA sampling.

Statistical Analysis

Cox proportional hazards survival model was used to estimate the risk of breast cancer incidence according to life events. The number of life events was modelled both as a continuous variable and as an ordinal variable in six categories (0-1, 2-3, 4-5, 6-7, 8- 9 and 10 or more) and estimated the risk for each category to allow for non-linear effects from less to more life events. Cox proportional hazards models were analysed adjusting for age (as the time-scale variable), with follow-up from birth, but entry into the model at the time of response to the 1981 questionnaire, until death, emigration or end-of-follow-up on December 31, 2018. Initial models were adjusted for age and sex, while multivariable models were additionally adjusted for potential breast cancer covariates listed above. Results from the models are reported with all covariates included, and the effect of each of the covariates (other than age or sex) on the risk of breast cancer was also examined (data not shown).

Within-pair Cox-regression models of life events and mortality were performed with the baseline hazard specific to each pair. The models yield overall estimates adjusted for familial and genetic factors shared by the twins in a pair. The proportional hazards assumption of the Cox regression model was examined by graphical methods of plotting log–log survival curves and tested using a global test based on Schoenfeld residuals.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Statistical analyses were done using Stata (v18, Statacorp, College Station, TX, USA). After obtaining our primary results, we ran sensitivity analyses examining the effect of birth cohort, in which our focus was on the effect sizes and not statistical significance testing. We report the means and standard deviations of continuously distributed covariates and the distributions of categorical covariates.

Linear models were used to test for association between breast cancer PRS and life event scores, while Cox regression models were used to examine the association between breast cancer PRS and breast cancer among all women with genetic data, and within DZ pairs discordant for breast cancer.

Linear models were employed to analyse the association between within pair difference in DNAm at 212 selected CpG sites [17] and the within pair difference in life event exposure. Two models were used, one using all available twin pairs and the other examining MZ pairs alone to fully control for genetic effects. Further, we examined whether the effect of stressful life events on DNAm aligned with the direction of DNAm associated with environmental BC risk as reported in Bode et al., 2024 [17]. Each CpG site with the same effect direction as in Bode et al. (2024) [17] and $p<0.05$ were considered as validated.

RESULTS

Stressful Life Events and Breast Cancer Incidence

Cohort analyses

During the 36-year follow-up, 719 incident breast cancer cases were recorded in our cohort. Participating women were aged 24 years and older in 1981 with a mean age of 39 (SD 12) years; 32% were MZ twins. The distribution of baseline characteristics in cases and controls is detailed in Supplemental Table 1.

The relationship between the number of stressful life events and breast cancer risk is shown in Table 1, whereby accumulating numbers of stressful life events increases the risk of breast cancer incidence. After adjusting for age, the hazard ratio for breast cancer per one-event increase in the total number of life events was 1.05 (95%CI 1.02-1.08) and remained consistent

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

at 1.04 (95%CI 1.00-1.07) after adjusting for additional potential confounders (zygosity, marital status, oral contraceptive use, social class, reproductive class, BMI, alcohol consumption, smoking, and leisure time physical activity). Being exposed to 10 or more stressful life events increased the risk of breast cancer incidence by 93% (Table 1). Within-pair analysis showed that for every additional stressful life event, MZ pairs had a higher within-pair HR than DZ pairs (1.14 [95%CI 1.01-1.29] vs 1.03 [95%CI 0.95-1.11]) for breast cancer, indicating a significant risk increase even after controlling for genetic background. The individual based analyses resulted in HRs with the same magnitude as the within DZ pair analysis (Table 1). The impact of known and proposed breast cancer risk factors on HRs for breast cancer incidence are detailed in Supplemental Table 2.

Stratification by birth decade

The extensive range of birth years in the cohort enabled the implementation of a segmented analysis to examine the impact of stressful life events across birth decades. Stratifying by birth decade revealed a modest increase in HR for those born before 1950 compared to those born after 1950 (HR 1.07 vs.1.01, Table 2, Supplemental Table 3). A deeper look into the other breast cancer risk factors uncovered some additional temporal differences: The most notable difference was the HR associated with use of oral contraceptives in those born before 1950 (HR 1.28; 95%CI 1.03-1.59) being appreciably higher than that of those born after 1950 (HR 1.01; 95%CI 0.75-1.35). Of other factors, being single and alcohol consumption exhibited modest variations when segregated by birth cohort (Supplemental Table 3).

Genetic breast cancer risk, cancer incidence and association with life events

Among women with a breast cancer PRS available, there were 364 breast cancer cases, while 4237 women were unaffected by the end of the 36-year follow-up. Adjusting for underlying family structure, we observed a 3% increase in breast cancer risk per SD increase in breast cancer PRS, with a HR of 1.03 (95%CI 0.92-1.15). Among 176 DZ twin sister pairs discordant for breast cancer diagnosis, a pairwise survival model yielded an HR of 1.03 (95%CI 0.89-1.20). For all women with stressful life events and breast cancer PRS data available (n=3906), the correlation between stressful life events and breast cancer PRS was -0.01 (95%CI -0.04 - 0.02).

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Environmental Breast Cancer Risk related DNA methylation

To assess if environmental breast cancer risk related DNAm is associated with stressful life events, we followed up our previous findings on 212 such CpG sites (Bode et al. 2024, [17]) in the 319 pairs with stressful life event score and sampled for DNAm analysis, on average, 16 years prior to breast cancer diagnosis. Their mean age at the 1981 questionnaire study was 40.0 (SD 9.8) years, and mean age at DNA sampling was 63.2 (8.9) years; these ages did not differ by zygosity. The intraclass correlation for the stressful life events score was 0.44 (95%CI 0.31-0.55) in MZ pairs indicating substantial non-genetic influence.

Our analysis implicated 42 of these previously identified CpG sites, with significant associations between stressful life events exposure and breast cancer risk related DNAm. These CpG sites are in 32 known genes (Supplemental File 1). In total, 197/212 CpG sites displayed association trend with stressful life events in the same direction as with breast cancer risk reported by Bode et al. (2024)[17] (Supplemental File 1). This significant skew towards hypomethylation associated with stressful life events exposure (binomial test $p = 1.19*10^{-41}$ is in line with our previous findings that showed that hypomethylation associates with an exposure to environmental breast cancer risk factors (Bode et al. 2024, [17]) (Figure 1).

To further isolate the influence of germline genetics on the observed effects of stressful life events on DNAm, we repeated the analysis solely on the 174 MZ pairs across all 212 DNAm sites (Supplemental File 1). The effect sizes of the within pair difference in stressful life events on the within pair difference in DNAm at the aforementioned 42 CpG sites were 36% larger in MZ pairs compared to DZ pairs (paired t-test, $p = 0.012$). This suggests that the observed associations between exposure to stressful life events and DNAm are largely independent of germline genetic factors. Further, within the MZ twin pairs the effect direction in 193/212 CpG sites was in line with that observed in Bode et al. (2024) [17] for breast cancer risk (binomial-test $p = 1.9*10^{-37}$).

As a sensitivity test, we repeated the analyses in the Older Finnish Twin Cohort with 154 female twin pairs (96 MZ and 58 DZ pairs) who remained cancer-free until the end of the follow-up in 2018, to exclude the potential effect of future cancer diagnosis on the association between DNAm at the 212 CpG sites and stressful life events. A total of 174 of the 212 CpG sites

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

(binomial test, $p = 1.1*10^{-20}$) showed directional consistency with Bode et al. (2024) [17] (Supplemental Table 1). We further restricted the analysis to the 96 MZ twin pairs (Supplemental Table 1), leveraging their matched germline genetics despite the smaller sample size. Even within this subset, 182 out of 212 CpG sites showed directional consistency with Bode et al. (2024) [17] (binomial test, $p = 9.8*10^{-28}$). These findings strengthen the evidence that DNAm patterns associated with environmental breast cancer risk might also be linked to stressful life events exposure, independent of both germline genetics and future breast cancer diagnosis.

DISCUSSION

We conducted a genetically informed study on the impact of stressful life events on breast cancer incidence with 36 years of follow-up. Our results emphasise the long-term association between stressful life events and breast cancer incidence, while pointing towards the limited role of genetic risk, as assessed by PRS and within-pair analyses of MZ twin pairs. Instead, by replicating environmentally driven breast cancer risk -related DNA methylation patterns, our study suggests that DNA methylation could play a role in mediating the effects of stressful life events on breast cancer incidence. Intriguingly, our analysis also revealed a birth-decade difference in risk factor association, with the HR of stressful life events varying across generations, suggesting a stronger association for those born before than after 1950**.**

It is well established that the balance of neuroendocrine hormones in women is easily affected by psychological trauma [22], though the impact of stressful life events on health outcomes, especially cancer, is far from fully understood. Major traumatic events are potent sources of stress and more easily attributable to adverse health events later in life [23]. In contrast, daily stresses in people's lives may be far more insidious in their long-term biological consequences. A previous study utilising the Finnish Twin Cohort showed a heightened association between accumulation of stressful life events, in particular of those of greatest a-priori-interest, such as divorce or death of a loved one, and increased risk of breast cancer incidence during 15 years of follow-up [11]. Our current study extends the follow-up period to 36 years after baseline assessment (1981) and includes a greater number of incident breast cancer cases. Life events occurring in the five years prior to assessment are qualified as "stressful" and quantified based on the 21-question list described by Lillberg et al. (2003) [11], with questions ranging from more

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

commonplace stressors, such as acquiring a loan, to the more unexpected and traumatic end of the spectrum, such as a death of a loved one. The increase in statistical power, attributed to the longer follow-up time and, consequently, the number of breast cancer cases, reasserts previous findings and shows that accumulating multiple stressful life events, regardless of their type, significantly increase the risk of breast cancer incidence in the long-term. Leveraging twin pair analyses, we established that the heightened risk of breast cancer is attributable to factors that are not shared between the twins, i.e., the increased risk in breast cancer incidence following stressful life events is independent of genetic background. Prior research has indicated that associations of stressful life events with breast cancer risk are not consistently observed. There are limitations in these analyses, such as potential recall bias in retrospective studies and the challenge of drawing temporal conclusions from studies with varying follow-up periods, as shorter follow-up periods could limit the ability to capture the full effect of stressful life events on breast cancer development, which can take years to manifest.

A noteworthy feature of our dataset is that individuals were born across different decades, often prior to other studies referenced in the literature [11]. While no clear trend in genetic risk across these cohorts was visible when stratifying the data based on decade of birth, we did observe a higher HR for stressful life events in individuals born before 1950. The observed difference in the risk ratio associated with stressful life events may be reflective of the turbulent times in Finnish history, such as the 1930s depression, World War II and the immediate post-war period, when Finland was much less developed than it is currently. Only from the 1950s did economic and social development happen within a long period of peace. Another notable finding is that of oral contraceptive use, whereby women born prior to 1950 had 28% increase in hazard ratio for breast cancer as compared to virtually no effect among those born after 1950. This may be due to the changing doses of oral contraceptives, as pills transitioned from high-dose oestrogen formulations (50-100 mcg) in the 1960s to lower doses (10-20mcg).

The epidemiology of breast cancer in Finland is evolving. The cumulative hazard for those born after 1950 in the current cohort was shown to be significantly higher than those born earlier (Supplemental Figure 1), reflecting the general trend of increasing breast cancer incidence in the Finnish population (Finnish Cancer Registry Incidence Figure from www.cancer.fi). However, birth cohort analyses in the entire Finnish population based on the NordCan database suggest that the age-specific breast cancer incidence of women born from 1950 onwards is no longer

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

increasing (Supplemental Figure 2, Nordcan visualisation from https://nordcan.iarc.fr/). The same change can be seen for Sweden as well in the Nordcan data. Our stratified analysis delineates the importance of timeframe when considering risk associations with breast cancer. There is potentially a paradigm shift in the epidemiology of breast cancer indexed by a major change in incidence for women born in the post-WWII decade in at least two Nordic countries. Our suggestive findings for stressful life events point to greater resilience among women born in the 1950s. These intriguing but preliminary findings warrant further investigation to consolidate the reasons underlying these changes in the epidemiology of breast cancer.

We used genetic and epigenetic data to explore potential explanations for the association of stressful life events and incident breast cancer. Firstly, genetic breast cancer risk assessment revealed, at most, a weak association between the breast cancer PRS and breast cancer incidence in our cohort. The hazard ratios ranged between 1.01 and 1.23, which are smaller than those reported in previous study by Mars et al. (2020) (HR = 1.71; 1.68–1.75) [4]. However, the effect size of the used breast cancer PRS was comparable in other European populations [24]. As our cohort comes from the same population of Finns as in Mars et al. (2020) [4], the smaller effect size is probably an indicator of other differences between the FinnGen cohort used by Mars et al. (2020) [4] and our twin cohort. The potentially diluted observation between breast cancer PRS and breast cancer incidence may be due to the substantially longer follow-up period compared to previous studies, as environmental risk factors may exert a stronger influence over time, masking the genetic contribution. While our findings on breast cancer PRS showed only weak association with breast cancer incidence, it is important to acknowledge limitations of PRS itself. The current PRS model explains a relatively small proportion (less than 1%) of the total breast cancer risk [4]. Therefore, it may not fully capture the estimated 31% risk attributable to genetic factors based on twin studies [2].

We did not see an association between breast cancer PRS and stressful life events score, suggesting that the effect of life events on breast cancer risk is not mediated by shared genetic liability. Though life events are considered to be external to the persons affected by them, the response to such events is in part affected by biological factors and hence possibly by genetic effects [25–27]. Likewise, the subjective experience of stressful life events is variable, and thus reporting of them may vary.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

In contrast to the genetic findings, our epigenetic results demonstrate an association between exposure to life events and specific DNAm patterns. These same DNAm patterns are also linked to an increased risk of breast cancer, particularly to breast cancer risk coming from environmental exposures (Bode et al, 2024) [17]. However, to ensure that this association with stressful life events is not simply a consequence of an underlying undiagnosed cancer, we have shown in a separate dataset of healthy twin pairs that similar effects of stressful life events on the environmental breast cancer risk DNAm patterns are also present there. These results suggest that the association between stressful life events and DNAm is likely independent of a subsequent breast or other cancer diagnosis. Prior research supports this connection: studies have shown that stress exposure can modify DNAm patterns [28–32]. Additionally, stressful life events can disrupt a woman's endocrine hormone balances [18]. Changes in endocrine hormones, in turn, are known to influence DNAm [33]. This suggests that life events might leave epigenetic marks on DNA directly or indirectly via changes in endocrine hormone levels, which could potentially affect the function of breast cancer susceptibility genes and thereby alter biological pathways that promote breast cancer development. The observed DNAm changes appearing, on average, 16 years after life event exposure but up to 11 years before breast cancer diagnosis strengthens the potential link between life events, and breast cancer development via DNAm, as these DNAm patterns are present during a critical time window for breast cancer initiation. However, future research is crucial to establish a causal relationship between DNAm and breast cancer risk, as well as with other cancers, especially ones sharing similar risk factor profiles or same driver genes. This could involve investigating the functional consequences of these DNAm changes on relevant pathways. Additionally, studies comparing DNAm in breast tissue instead of blood samples would provide valuable insights.

Our study has some limitations: Our results are confined to women of European origin, specifically of Finnish descent, underscoring the necessity of conducting studies on diverse populations before making broader generalisations. We acknowledge that some of our analyses may lack statistical power due to limited sample size. Our study did not include data on high-impact breast cancer variants, such as those in *BRCA1* and *BRCA*2, limiting a comprehensive examination of their potential influence on breast cancer risk in the current study as these variants do interact with the PRS in a non-additive manner, as demonstrated by Mars et al.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

(2020) [4]. However, given the number of breast cancer cases, any analyses of high-impact breast cancer genes would be severely constrained.

CONCLUSION

This investigation contributes to our understanding of the delicate aetiology of breast cancer and further solidifies a concerning trend: stressful life events appear to be a major long-acting risk factor for breast cancer. Cumulative stressful life events emerged as a potent risk factor, with a hazard ratio exceeding that of other measured factors across generations, especially for women born prior to 1950. While genetics has an effect, it is likely weak, as our analysis suggests that the majority of breast cancer incidence is likely influenced by non-genetic factors. Additionally, epigenetic changes linked to breast cancer risk were associated with stress exposure. Remarkably, these changes persisted on average at least 16 years after the stressful life events and, on average, 11 years before diagnosis. Our findings highlight the critical need to understand the role of stress in breast cancer aetiology.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Declarations

Funding

This research was supported by funding from Biology of Trauma Initiative funding at the Broad Institute of MIT and Harvard (award 6910369-5500001940) to Jaakko Kaprio, Academy of Finland] (#328685, 307339, 297908 and 251316), Sigrid Juselius Foundation, Minerva Foundation and Medicinska Understödsföreningen Liv o Hälsa r.f. to Miina Ollikainen, Cancer Foundation Finland (#67-6796) and Yrjö Jahnsson Foundation (#20237645) to Hannes Bode, and Emil Aaltonen Foundation to Mikaela Hukkanen (#23005).

Competing Interests

The Authors have no competing interests to declare.

Author Contributions

The study was initiated and designed by Jaakko Kaprio, Elissar Azzi, Hannes Bode and Miina Ollikainen. Data acquisition and preparation were performed by Jaakko Kaprio, Miina Ollikainen, Hannes Bode, Teemu Palviainen and Mikaela Hukkanen. Analysis of the data was conducted by Jaakko Kaprio, Hannes Bode, and Elissar Azzi. The first draft of the manuscript was written by Elissar Azzi and Hannes Bode, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Ethics approval

This study has been approved by the appropriate national research ethics committees, with the most recent approval granted by the Hospital District of Helsinki and Uusimaa ethics board in 2018 (#1799/2017). Permission for linkage to the Finnish Cancer Registry was provided by the Finnish Social and Health Data Permit Authority Findata (THL/6353/14.06.00/2023).

Consent to participate

Blood samples for DNA analyses were collected from each participant after they had signed a written informed consent.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References:

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 71:209–249

2. Möller S, Mucci LA, Harris JR, et al (2016) The heritability of breast cancer among women in the nordic twin study of cancer. Cancer Epidemiol Biomarkers Prev 25:145–150

3. Turchiano A, Piglionica M, Martino S, et al (2023) Impact of High-to-Moderate Penetrance Genes on Genetic Testing: Looking over Breast Cancer. Genes (Basel). https://doi.org/10.3390/GENES14081530

4. Mars N, Widén E, Kerminen S, et al (2020) The role of polygenic risk and susceptibility genes in breast cancer over the course of life. Nat Commun. https://doi.org/10.1038/S41467-020-19966-5

5. Britt KL, Cuzick J, Phillips KA (2020) Key steps for effective breast cancer prevention. Nat Rev Cancer 2020 208 20:417–436

6. Antonova L, Aronson K, Mueller CR (2011) Stress and breast cancer: from epidemiology to molecular biology. Breast Cancer Res 13:208

7. Ma W, Liu P, Zheng J, et al (2021) Immune and nonimmune mechanisms mediate the mental stress-induced tumor growth in a xenograft model of breast cancer. Cell Death Dis. https://doi.org/10.1038/S41419-021-04280-9

8. Gonggrijp BMA, Van De Weijer SGA, Bijleveld CCJH, Boomsma DI, Van Dongen & J, Nl BMAG (2024) Negative Life Events and Epigenetic Ageing: a Study in the Netherlands Twin Register. bioRxiv 2024.02.20.581138

9. Schoemaker MJ, Jones ME, Wright LB, Griffin J, McFadden E, Ashworth A, Swerdlow AJ (2016) Psychological stress, adverse life events and breast cancer incidence: a cohort investigation in 106,000 women in the United Kingdom. Breast Cancer Res. https://doi.org/10.1186/S13058-016-0733-1

10. Bahri N, Fathi Najafi T, Homaei Shandiz F, Tohidinik HR, Khajavi A (2019) The relation between stressful life events and breast cancer: a systematic review and meta-analysis of cohort studies. Breast Cancer Res Treat 176:53–61

11. Lillberg K, Verkasalo PK, Kaprio J, Teppo L, Helenius H, Koskenvuo M (2003) Stressful life events and risk of breast cancer in 10,808 women: a cohort study. Am J Epidemiol 157:415–423

12. Penner-Goeke S, Bothe M, Rek N, et al (2023) High-throughput screening of glucocorticoid-induced enhancer activity reveals mechanisms of stress-related psychiatric disorders. Proc Natl Acad Sci U S A 120:e2305773120

13. Ising M, Holsboer F (2006) Genetics of stress response and stress-related disorders. Dialogues Clin Neurosci 8:433–444

14. Jablonka EVA, Raz GAL (2009) Transgenerational epigenetic inheritance: prevalence,

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

mechanisms, and implications for the study of heredity and evolution. Q Rev Biol 84:131–176

15. Li M, D'Arcy C, Li X, Zhang T, Joober R, Meng X (2019) What do DNA methylation studies tell us about depression? A systematic review. Transl Psychiatry. https://doi.org/10.1038/S41398-019-0412-Y

16. Parade SH, Huffhines L, Daniels TE, Stroud LR, Nugent NR, Tyrka AR (2021) A systematic review of childhood maltreatment and DNA methylation: candidate gene and epigenome-wide approaches. Transl Psychiatry. https://doi.org/10.1038/S41398-021-01207-Y

17. Bode, Hannes Frederik A, He L, Hjelmborg J, Kaprio J, Ollikainen M (2024) Pre-diagnosis blood DNA methylation profiling of twin pairs discordant for breast cancer points to the importance of environmental risk. medRxiv. https://doi.org/10.1101/2023.08.15.23293985

18. Kaprio J, Bollepalli S, Buchwald J, et al (2019) The Older Finnish Twin Cohort - 45 Years of Follow-up. Twin Res Hum Genet 22:240–254

19. Hukkanen M, Kankaanpää A, Heikkinen A, Kaprio J, Cristofari R, Ollikainen M (2024) Human reproduction comes at the expense of faster aging and a shorter life. bioRxiv 2024.07.18.603826

20. Michailidou K, Lindström S, Dennis J, et al (2017) Association analysis identifies 65 new breast cancer risk loci. Nature 551:92–94

21. Gustavson K, Torvik FA, Davey Smith G, Røysamb E, Eilertsen EM (2024) Familial confounding or measurement error? How to interpret findings from sibling and co-twin control studies. Eur J Epidemiol 39:587–603

22. Ravi M, Stevens JS, Michopoulos V (2019) Neuroendocrine pathways underlying risk and resilience to PTSD in women. Front Neuroendocrinol 55:100790

23. Cohen S, Murphy MLM, Prather AA (2019) Ten Surprising Facts About Stressful Life Events and Disease Risk. Annu Rev Psychol 70:577

24. Mars N, Kerminen S, Feng YCA, et al (2022) Genome-wide risk prediction of common diseases across ancestries in one million people. Cell genomics. https://doi.org/10.1016/J.XGEN.2022.100118

25. Thapar A, Mcguffin P (1996) Genetic influences on life events in childhood. Psychol Med 26:813–820

26. Kendler KS, Karkowski LM, Prescott CA (1999) The assessment of dependence in the study of stressful life events: validation using a twin design. Psychol Med 29:1455–1460

27. Stein MB, Jang KL, Taylor S, Vernon PA, Livesley WJ (2002) Genetic and environmental influences on trauma exposure and posttraumatic stress disorder symptoms: a twin study. Am J Psychiatry 159:1675–1681

28. Dunn EC, Soare TW, Zhu Y, Simpkin AJ, Suderman MJ, Klengel T, Smith ADAC, Ressler KJ, Relton CL (2019) Sensitive Periods for the Effect of Childhood Adversity on

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

DNA Methylation: Results From a Prospective, Longitudinal Study. Biol Psychiatry 85:838–849

29. Yan Q, Forno E, Cardenas A, et al (2021) Exposure to violence, chronic stress, nasal DNA methylation, and atopic asthma in children. Pediatr Pulmonol 56:1896

30. Smith AK, Ratanatharathorn A, Maihofer AX, et al (2020) Epigenome-wide meta-analysis of PTSD across 10 military and civilian cohorts identifies methylation changes in AHRR. Nat Commun. https://doi.org/10.1038/S41467-020-19615-X

31. Scorza P, Duarte CS, Lee S, Wu H, Posner J, Baccarelli A, Monk C (2023) Stage 2 Registered Report: Epigenetic Intergenerational Transmission: Mothers' Adverse Childhood Experiences and DNA Methylation. J Am Acad Child Adolesc Psychiatry 62:1110–1122

32. Kuan PF, Waszczuk MA, Kotov R, Marsit CJ, Guffanti G, Gonzalez A, Yang X, Koenen K, Bromet E, Luft BJ (2017) An epigenome-wide DNA methylation study of PTSD and depression in World Trade Center responders. Transl Psychiatry 7:e1158

33. García-Carpizo V, Ruiz-Llorente L, Fraga M, Aranda A (2011) The growing role of gene methylation on endocrine function. J Mol Endocrinol. https://doi.org/10.1530/JME-11-0059

34. Kujala UM, Palviainen T, Pesonen P, et al (2020) Polygenic Risk Scores and Physical Activity. Med Sci Sports Exerc 52:1518

35. Therneau T (2023) A Package for Survival Analysis in R.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Tables and Figures

Table 1

Survival model estimates of hazard ratios associated with the number of stressful life events and every one event increase in the sum of reported life events in 1981 and breast cancer incidence between 1982 and 2018 in all individuals and within MZ and DZ pairs.

Number of events	Age-adjusted HR (95%CI)	Fully-adjusted HR (95%CI)	Within DZ pairs HR (95%CI)	Within MZ pairs HR (95%CI)
0-1	1	1	1	1
2-3	1.3 (1.04-1.62)	1.24 (1.00-1.54)	1.08 (0.65-1.80)	1.97 (0.85-4.57)
4-5	1.45 (1.15-1.81)	1.33 (1.05-1.66)	1.14 (0.65-2.00)	2.2 (0.97-5.00)
6-7	1.35 (1.04-1.75)	1.19 (0.91-1.56)	1.05 (0.57-1.94)	1.81 (0.68-4.87)
8-9	1.41 (0.98-2.03)	1.24 (0.85-1.80)	1.08 (0.49-2.38)	1.91 (0.52-7.06)
10-17	1.93 (1.19-3.14)	1.7 (1.04-2.78)	1.3 (0.36-4.68)	7.15 (1.26-40.68)
Sum of life events – increase in HR per one additional event	1.05 (1.02-1.08)	1.04 (1.00-1.07)	1.03 (0.95-1.11)	1.14 (1.01-1.29)

Multivariable model adjusts for age, zygosity, marital status, oral contraceptive use, social class, reproductive class, BMI, alcohol consumption, smoking, leisure time physical activity. Results are shown by categories of the number of life events and for the continuous variable of a total sum score of life events. HR, hazard ratio; CI confidence interval

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2

Hazard ratios for sum of life events by birth cohort segments

Birth cohort	All	Before 1940	1940-1949	1950-1957
Age at response to 1981 survey (min, max)		41.8, 92.2	31.8, 43.2	23.8, 33.4
Sum of life events – increase in HR per one additional event (95%CI)	1.05 (1.02-1.08)	1.04 (0.98-1.11)	1.07 (1.01-1.12)	1.01 (0.96-1.06)
Number of breast cancer cases	719	222	230	267
Number of women	10342	3349	3049	3944

HR, Hazard ratio; CI, Confidence interval

medRxiv preprint doi: https://doi.org/10.1101/2024.08.27.24312571; this version posted August 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

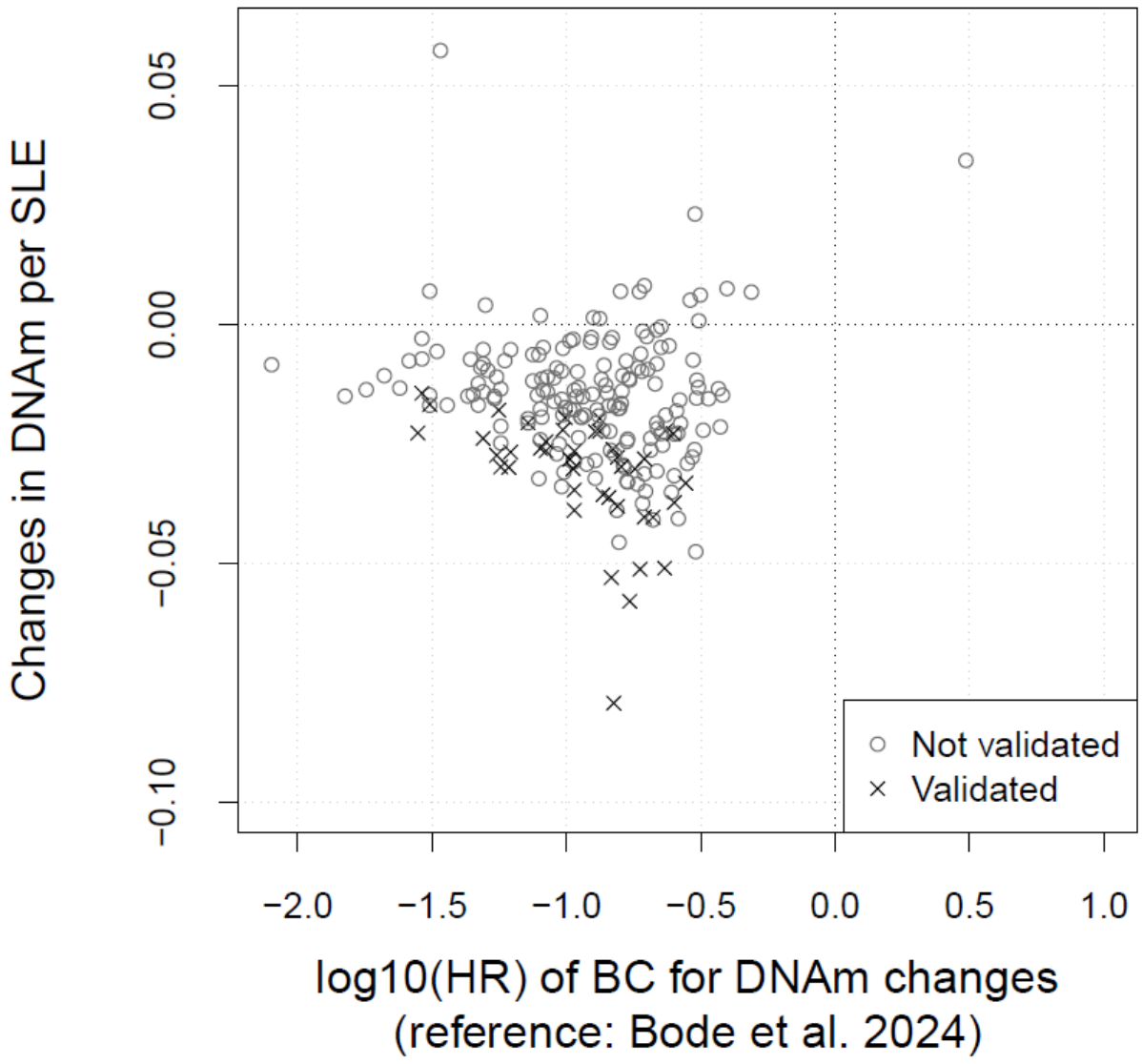

Figure 1: Association between stressful life events exposure and DNAm differences among all pairs

On the y-axis the effects of exposure to stressful life events (SLE) on breast cancer related differences in DNAm in twin pairs (172 DZ and 147 MZ pairs) for 212 CpG sites identified in Bode et al. (2024) [17]. On the x-axis the hazard ratios (HR) of breast cancer (BC) in relation to changes in DNAm according to Bode et al. (2024) [17]. The validated data points indicate that the association between stressful life events and DNAm have a p-value < 0.05.

➔ This Figure was created with R version 4.3.2.