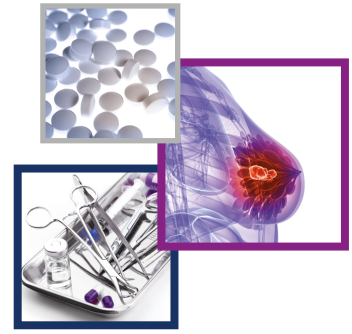
Short Communication

For reprint orders, please contact: reprints@futuremedicine.com

Breast Cancer Management



Circulating sphingomyelins on estrogen receptor-positive and estrogen receptor-negative breast cancer-specific survival

Charleen D Adams*,[1]

[1]City of Hope, Beckman Research Institute, 1500 E. Duarte Road, Duarte, CA 91010, USA

*Author for correspondence: chaadams@coh.org

Aim: This study aims to determine whether a causal relationship exists between circulating sphingomyelins and breast cancer-specific survival, since, if one does, sphingomyelins could be studied as a therapeutic target in the management of breast cancer. **Patients/materials & methods:** Mendelian randomization is used here to investigate whether higher levels of circulating sphingomyelins impact breast cancer-specific survival for estrogen receptor-negative (ER–) and estrogen receptor-positive (ER+) patients. **Results:** The results suggest a null effect of sphingomyelins for ER– breast cancer-specific survival and a protective effect for ER+ breast cancer-specific survival. Sensitivity analyses implicate low-density lipoprotein cholesterol as a potential confounder. **Conclusion:** Future studies should replicate and triangulate the present findings with other methods and tease out the roles of sphingomyelins and low-density lipoprotein cholesterol on breast cancer-specific survival.

First draft submitted: 13 January 2020; Accepted for publication: 18 May 2020; Published online: 26 June 2020

Keywords: breast cancer survival • lipids • mendelian randomization • metabolism • sphingomyelins

Breast cancer is the leading cause of cancer-related death for women aged 20–59 years in USA. More than 40,000 breast cancer patients died in 2018 [1]. A deeper understanding of the molecular mechanisms leading to breast cancer progression – processes that impact cell growth, invasion, angiogenesis, metastasis and, ultimately, survival – is, therefore, urgent to prevent deaths caused by this common cancer. Mechanistic insights could lead to new drug targets and interventions to better manage the disease.

To that aim, sphingolipids are key regulatory molecules that control aspects of cell growth arrest and proliferation. They are also essential for cellular structural integrity and have been studied to some extent for their role in cancer progression [2]. Notwithstanding this, proteins – not lipids – have had the lion's share of attention as signaling molecules in cancer-progression research [2].

Various sphingolipids impinge on a major pathway downstream of *HER2* [3] – PI3K/AKT – which represses apoptosis and autophagy and is involved in cancer progression and drug resistance [4]. Somatic mutations in *PIK3CA* occur in 25–35% of breast cancers, with enrichment in *HER2* and hormone receptor-positive tumors [5]. Thus, sphingolipids are important, potentially modifiable targets to prevent breast cancer progression. A few examples of sphingolipids include ceramide, sphingosine, sphingosine-1-phosphate and sphingomyelin.

Briefly, ceramide is a tumor-suppressor lipid involved in cell growth arrest and induction of cell death [6]. Intracellular ceramides are released by apoptotic cells and, thus, cell death within the tumor is a source of ceramides that belongs to the tumor, not the host [7]. Higher levels of ceramides have been associated with less tumor aggressiveness in breast cancer patients [8]. Conversely, lower levels of endogenous ceramides are involved in multidrug resistance [9].

Ceramide can serve as a structural and metabolic precursor for sphingomyelin [9]. Sphingomyelin, which can be reciprocally hydrolyzed to ceramide [9], has been reported to be higher in breast cancer than normal breast tissue [2]. While sphingomyelin is the major sphingolipid in cell membranes and has been investigated for its chemopreventive

Future Medicine

and chemotherapeutic potential [6], the role of circulating levels of sphingomyelin on breast cancer-specific survival is currently unknown. If a causal relationship exists between circulating sphingomyelins and breast cancer-specific survival, this knowledge could be exploited for the development of new drug targets and/or other interventions to improve survival for breast cancer patients.

Here, circulating levels of sphingomyelin are explored for their role in breast cancer-specific survival. Specifically, Mendelian randomization (MR) is used to investigate whether higher levels of circulating sphingomyelins impact breast cancer-specific survival for estrogen receptor-negative (ER-) and estrogen receptor-positive (ER+) patients.

What is MR?

MR is an analytic technique that uses genetic variants as instruments to examine the effect of an exposure (i.e., independent variable; in the present study sphingomyelins) on an outcome (i.e. dependent variable; in the present study, ER– and ER+ breast cancer survival).

Though MR is a methodological strategy that has been around since the early 1990s [10] and is increasingly commonly used with the wide availability of large genome-wide association (GWA) studies, researchers who have not performed an MR analysis may be unfamiliar with it.

MR is referred to as a 'causal' method, in contrast to observational studies for which randomization is not done, usually for ethical reasons. Because its strategy is paradigm shifting, it has been referred to as a 'gene-based hack' by a freelance journalist writing for *Nature* [11]. In MR, quasi-randomization happens on genotype, 'hacking' the following natural processes: Mendel's laws of inheritance and the temporal assignment of genotype at conception. Exploiting these means that, at the population level, genetic variants strongly associated with an exposure of interest are generally, though not always, independent of confounders that can distort observational estimates [12,13]. An exception might be population stratification if a family-based design is not used. Since genotype assignment occurs at conception, capitalizing on this overcomes most instances of reverse causation.

MR has certain important caveats, assumptions that must hold in order for its results to be valid. First, the genetic variants (usually single-nucleotide polymorphisms [SNPs]) serving as proxies for the exposure trait of interest (e.g., circulating sphingomyelins) must be robustly associated with the exposure trait (conventionally SNPs associated at genome-wide significance are used satisfy the robustness assumption, $p < 5 \times 10^{-8}$). Second, the genetic variants serving as exposure-trait proxies must not be associated with confounders of the exposure and outcome traits of interest. Third, the genetic variants serving as proxies must not be directly associated with the outcome of interest through a pathway other than the exposure of interest. This third assumption is what is meant by 'no pleiotropy'. MR sensitivity estimators have been developed to examine violations to the no-pleiotropy assumption.

Originally MR was performed within a single sample. With the availability of massive GWA studies, the MR procedure was adapted to use summary data from two GWA studies: one for the exposure and one for the outcome. This is known as two-sample MR and is the method used for the present study of circulating sphingomyelins on breast cancer-specific survival.

Materials & methods

Two-sample MR data sources

Sample 1: the sphingomyelin summary genetic data come from Kettunen *et al.* (2016), which performed a GWA study of 123 circulating metabolites – including sphingomyelins – in European participants (n = 13,476 for sphingomyelins) [14]. From this, independent (those not in linkage disequilibrium; $R^2 < 0.01$) SNPs associated at genome-wide significance ($p < 5 \times 10^{-8}$) with a standard-deviation (SD) increase in circulating sphingomyelins were identified and the summary statistics extracted. Nine SNPs were available for this purpose. The summary data for the Kettunen GWA study are publicly available through MR-Base [15], a platform that contains many GWA studies whose summary statistic data have been made public for use in two-sample MR [16].

Sample 2: the summary genetic data for ER− and ER+ breast cancer-specific survival come from two publicly available GWA studies of breast cancer-specific survival: one for ER− and one for ER+ breast cancer-specific survival. These were performed by Guo *et al.* (2015) [17]. Guo *et al.* pooled data from multiple breast cancer case cohorts of European ancestry, comprising eight datasets. ER status was obtained from medical records, followed up by immunohistochemistry on either tumor tissue microarrays or whole-section slides. Cox proportional hazards models were fitted to assess the relationship between genotype and breast cancer-specific mortality. The survival analysis for ER− contained 6,881 patients with ER− breast cancer, of which 920 died. The GWA study for ER+

Table 1. Summary statistic data for two-sample Mendelian randomization of circulating sphingomyelins on ER– and ER+ breast cancer-specific survival.

SNP	Effect allele	Other allele	EAF exposure	β exposure	SE exposure	p exposure	EAF outcome	Log HR outcome	SE outcome	p outcome
ER–										
rs10402112	A	T	0.10	-0.14	0.02	1.34E-11	0.10	-0.025	0.09	0.7874
rs11591147	T	G	0.03	-0.37	0.04	3.74E-20	0.01	0.087	0.27	0.7471
rs13392272	T	C	0.40	0.07	0.01	2.08E-08	0.48	0.005	0.05	0.9274
rs141622900	A	G	0.03	-0.29	0.04	5.75E-13	0.05	0.162	0.12	0.1704
rs174418	C	T	0.56	-0.08	0.01	1.09E-10	0.59	-0.033	0.05	0.5040
rs1800588	T	C	0.25	0.12	0.01	3.20E-16	0.22	0.053	0.06	0.3547
rs190934192	A	G	0.03	-0.24	0.04	4.47E-08	0.02	0.181	0.27	0.5099
rs629301	T	G	0.79	0.09	0.01	3.13E-09	0.76	−0.122	0.06	0.0288
rs75679663	A	C	0.03	-0.41	0.04	1.95E-21	0.01	-0.106	0.30	0.7241
ER+										
rs10402112	A	T	0.10	-0.14	0.02	1.34E-11	0.10	0.015	0.08	0.8425
rs11591147	T	G	0.03	-0.37	0.04	3.74E-20	0.01	0.042	0.21	0.8416
rs13392272	T	C	0.40	0.07	0.01	2.08E-08	0.47	0.024	0.04	0.5724
rs141622900	A	G	0.03	-0.29	0.04	5.75E-13	0.05	0.342	0.09	0.0002
rs174418	C	T	0.56	-0.08	0.01	1.09E-10	0.58	0.009	0.04	0.8203
rs1800588	T	C	0.25	0.12	0.01	3.20E-16	0.23	-0.048	0.05	0.3241
rs190934192	A	G	0.03	-0.24	0.04	4.47E-08	0.02	-0.234	0.21	0.2553
rs629301	T	G	0.79	0.09	0.01	3.13E-09	0.78	-0.043	0.05	0.3673
rs75679663	A	C	0.03	-0.41	0.04	1.95E-21	0.01	0.197	0.22	0.3599

Summary data for sphingomyelin-associated SNPs ('exposure') and the sphingomyelin-associated SNPs extracted from the ER– and ER+ breast cancer-specific survival genome-wide association studies ('outcome').
β: Beta estimate; EAF: Effect allele frequency; Log HR: Log hazards ratio; SE: Standard error; SNP: Single-nucleotide polymorphism.

survival contained 23,059 patients with ER+ breast cancer, of which 1333 died. The summary data for these breast cancer-specific survival GWA studies are publicly available through MR-Base.

The summary statistics for the nine sphingomyelin SNPs are displayed in Table 1.

Statistical approach

Estimates of the proportion of variance in circulating sphingomyelins explained by the nine-SNP sphingomyelins instrument (R^2) and the strength instrument (F-statistic) were generated (conventionally F-statistics < 10 are weak). A range of power estimates were also calculated for both ER– and ER+ breast cancer-specific survival using the 'mRnd' MR power calculator [18,19] (Figure 1).

The log-hazards ratios for breast cancer-specific survival per SD-increase in circulating sphingomyelins were calculated, using the inverse-variance weighted (IVW) MR method. The 'TwoSampleMR' package [16] was used for the MR analysis. The results were exponentiated to hazards ratios (HR) for interpretability.

Sensitivity analyses

Statistically based sensitivity estimators that make different assumptions about the underlying nature of pleiotropy can be used to appraise pleiotropic bias in MR studies. Three were chosen to complement the primary IVW causal tests: MR-Egger regression including its intercept test, weighted median, and weighted mode estimations. If there is a major violation to the MR assumption about no pleiotropy, we would expect that the effect estimates for the various MR estimators to differ in their magnitudes or directions. This comparison with the IVW method is how these estimators are used in the present study (for a detailed explanation of the different assumptions made by the various MR estimators, see the supplement to Yarmolinsky *et al.* (2019) [20, 21]).

In addition, statistical tests for heterogeneity were performed, since variability in the causal estimates between SNPs can indicate pleiotropy. The test for heterogeneity was performed using Cochrane's Q-statistic ($p < 0.05$ indicates a lack of heterogeneity) [20].

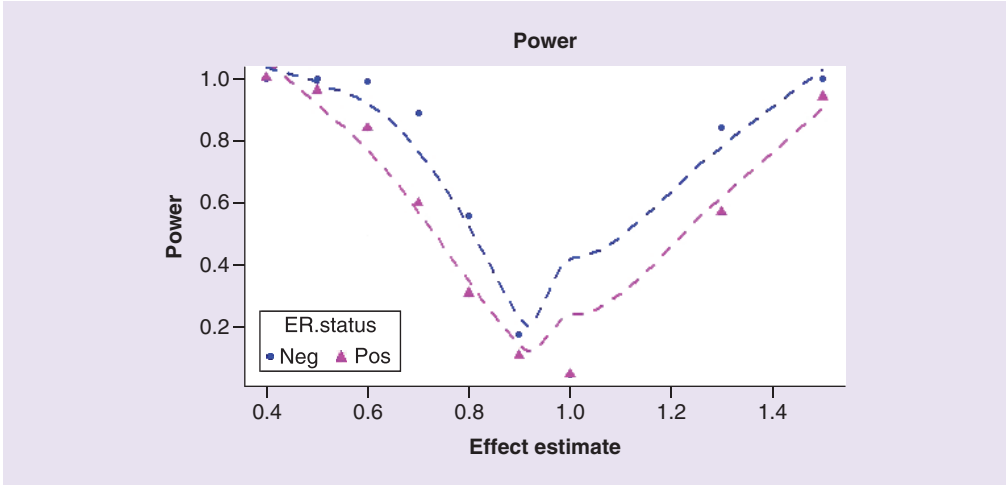


Figure 1. Range of power estimates for the Mendelian randomization analyses of sphingomyelins on ER– (neg) and ER+ (pos) breast cancer-specific survival.

Beyond the formal statistical tests for pleiotropy, the nine SNPs instrumenting circulating sphingomyelins were looked up in PhenoScanner, a database containing known phenotype-genotype associations [22,23]. The look-up in PhenoScanner provides knowledge about potential pleiotropic pathways. The sphingomyelin SNPs were associated with various cholesterols, including low-density (LDL) cholesterol and high-density (HDL) cholesterol. This was expected, since sphingomyelin is a sphingolipid in peripheral blood [24–27], and the metabolisms of sphingomyelin and cholesterol are partly coordinated [28].

MR sensitivity analyses of the influence of LDL cholesterol and HDL cholesterol on sphingomyelins were performed to demonstrate the close causal relationships between these molecules (Supplementary Tables 1 & 2 contain the SNP characteristics for the MR tests of LDL cholesterol on sphingomyelins and HDL cholesterol on sphingomyelins, respectively). The GWA data sources for the LDL cholesterol and HDL cholesterol instruments were obtained from Willer *et al.* (2013) and are available through MR-Base. For these tests, MR radial regression was used to detect and remove SNP outliers [29]. As with the primary tests for sphingomyelins on ER– and ER+ breast cancer-specific mortality, MR-Egger regression, weighted median and weighted mode estimators were also calculated.

Finally, a sensitivity analysis for LDL cholesterol and HDL cholesterol in relation to ER+ breast cancer-specific survival was done. (Since the main results for ER– breast cancer-specific survival were null, LDL cholesterol and HDL cholesterol were not examined in relation to ER– breast cancer-specific survival.) As mentioned above, the instruments for the LDL cholesterol and HDL cholesterol instruments were obtained from Willer *et al.* (2013) GWA study. Supplementary Tables 3 and 4 contain the SNP characteristics for LDL cholesterol and HDL cholesterol, respectively, and the summary data for these same SNPs as extracted from the GWA study for ER+ breast cancer-specific survival. The study population included up to 188,578 participants of European ancestry from 37 cohorts. For most, blood lipids were measured after 8 h of fasting. Those on lipid-lowering medications were excluded. The GWA study was adjusted for age and sex [30]. The Willer *et al.* (2013) GWA studies of LDL cholesterol and HDL cholesterol were selected because of the large sample sizes, which maximizes power to detect effects.

The associations between LDL cholesterol and HDL cholesterol on sphingomyelins and LDL cholesterol and HDL cholesterol on ER+ breast cancer-specific survival were done to assess confounding (see Discussion).

All described analyses were performed in R version 3.5.2.

Results

The instrument for sphingomyelins had an R^2 = 0.04 and F-statistic of 55. There was approximately 80% or more power to detect true effects between 0.4 and 0.6 for both ER– and ER+ breast cancer-specific survival (Figure 1).

Table 2. Mendelian randomization and sensitivity estimator results for the tests of circulating sphingomyelin levels on ER– and ER+ breast cancer-specific survival.

Method	SNPs	HR	Lower 95% CI	Upper 95% CI	p-value	Q	Q p-value
ER– breast cancer-specific survival							
Inverse variance weighted	9	0.86	0.58	1.27	0.44	8	0.42
MR Egger	9	0.73	0.31	1.74	0.50	8	0.34
Weighted median	9	1.08	0.61	1.91	0.78		NA
Weighted mode	9	1.33	0.53	3.34	0.57		NA
ER+ breast cancer-specific survival							
Inverse variance weighted	9	0.65	0.45	0.94	0.02	11	0.20
MR Egger	9	0.41	0.21	0.83	0.04	8	0.30
Weighted median	9	0.67	0.42	1.06	0.08		NA
Weighted mode	9	0.74	0.36	1.52	0.44		NA

HR: Hazards ratio; NA: Not applicable; Q: Cochrane's Q-statistic; SNP: Single-nucleotide polymorphism.

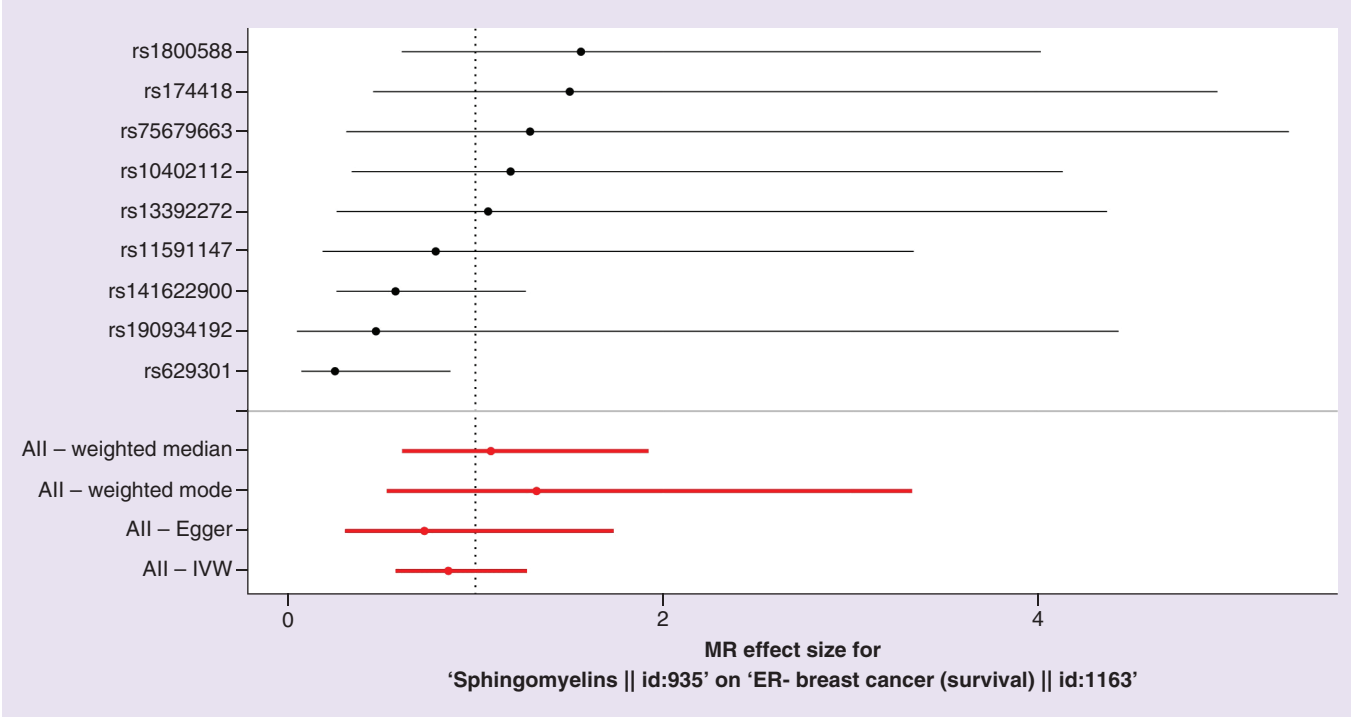


Figure 2. Causal estimates of individual single-nucleotide polymorphisms and meta-analyzed estimates for the effect of circulating sphingomyelins on ER– breast cancer-specific survival. The x-axis displays the hazards ratios and confidence intervals. The dotted line represents 1. The confidence intervals cross 1 for all the meta-analytic MR estimators (in red), indicating the relationship is not causal. IVW: Inverse-variance weighted; MR: Mendelian randomization.

ER– results

The MR analysis of circulating sphingomyelins on ER– breast cancer-specific survival was null: IVW estimate (HR) per 1-SD increase in circulating sphingomyelins = 0.86; 95% CI: 0.58, 1.27; p = 0.44) (Table 2 & Figure 2). Though the MR-intercept test (intercept = 1.02; 95% CI: 0.92, 1.14; p = 0.69) and test for heterogeneity with Cochrane's Q-statistic demonstrated no evidence for pleiotropy, the MR sensitivity estimators ranged in both their magnitudes and directions (Table 2). Overall, there was no evidence for an impact of sphingomyelins on ER– breast cancer-specific survival.

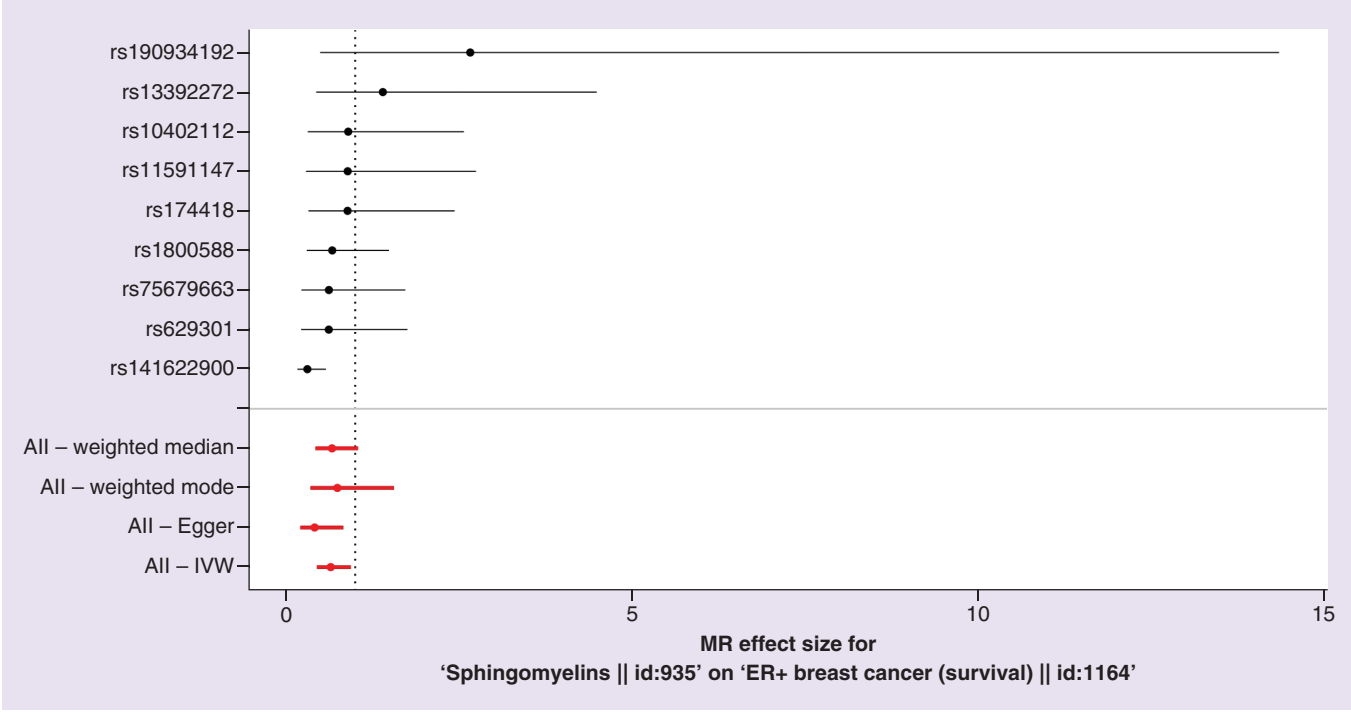


Figure 3. Causal estimates of individual SNPs and meta-analyzed estimates for the effect of circulating sphingomyelins on ER+ breast cancer-specific survival. The x-axis displays the hazards ratios and confidence intervals. The dotted vertical line represents 1. When the confidence intervals do not cross 1, as is the case for the inverse-variance weighted and MR-Egger estimates, this indicates a causal association.
IVW: Inverse-variance weighted; MR: Mendelian randomization.

ER+ results

The MR analysis of circulating sphingomyelins on ER+ breast cancer-specific survival indicated a causal, protective relationship: IVW estimate (HR) per 1-SD increase in circulating sphingomyelins = 0.65; 95% CI: 0.45, 0.94; p = 0.02) (Table 2 & Figure 3). The MR intercept (intercept = 1.07; 95% CI: 0.98, 1.16; p = 0.19) and test for heterogeneity with Cochrane's *Q*-statistic demonstrated no evidence for pleiotropy. There was consistency in the directions of the estimates across the MR sensitivity estimators and a dip in the magnitude for the MR-Egger estimate (MR-Egger estimate = 0.41; 95% CI: 0.21, 0.83; p = 0.04) in comparison with the IVW estimate.

PhenoScanner-based sensitivity analyses

The MR analyses of LDL cholesterol and HDL cholesterol on circulating sphingomyelins both suggested causal relationships (Table 3). The evidence for the impact of LDL cholesterol on circulating sphingomyelins was strong: a 1-SD increase in circulating LDL cholesterol led to an increase in circulating sphingomyelins: IVW (β) estimate = 0.52; 95% CI: 0.46, 0.58; p = 3.74E-59. The sensitivity estimators were comparable in both direction and magnitude. (Notice that the results are β estimates and not HRs). There was no evidence of heterogeneity or pleiotropy from either from the Cochrane's Q-statistic test or the MR-Egger intercept test (estimate = -0.001; 95% CI: -0.007, 0.005; p = 0.74).

The evidence was strong for HDL cholesterol, as well, though the IVW estimate and the MR-Egger estimate differed in magnitude (Table 3). A 1-SD increase in HDL cholesterol led to an increase in circulating sphingomyelins: IVW (β) estimate 0.26; 95% CI: 0.18, 0.34; p = 4.26E-10; MR-Egger estimate = 0.13. There was no evidence of heterogeneity from the Cochran's Q-statistic test, but some evidence of pleiotropy MR-Egger intercept test (estimate = 0.006; 95% CI: 0.00001; 0.01; p = 0.05).

The MR analysis of LDL cholesterol on ER+ breast cancer-specific survival indicated a causal, protective relationship: IVW estimate (HR) per 1-SD increase in LDL cholesterol = 0.81; 95% CI: 0.67, 0.97; p = 0.02) (Table 4). The MR-intercept (intercept = 1.00; 95% CI: 0.90, 1.02; p = 0.72) and test for heterogeneity with

Table 3. Causal estimates for LDL cholesterol and HDL cholesterol on circulating sphingomyelins.

Method	SNPs	β	Lower 95% CI	Upper 95 CI	p-value	Q	Q p-value
LDL cholesterol on circulating sphingomyelins							
Inverse variance weighted	67	0.52	0.46	0.58	3.74E-59	59	0.71
MR Egger	67	0.53	0.44	0.63	8.96E-16	59	0.68
Weighted median	67	0.56	0.46	0.66	3.58E-29	NA	NA
Weighted mode	67	0.56	0.47	0.66	3.24E-17	NA	NA
HDL cholesterol on circulating sphingomyelins							
Inverse variance weighted	69	0.26	0.18	0.34	4.26E-10	53	0.90
MR Egger	69	0.13	-0.02	0.28	0.09	49	0.95
Weighted median	69	0.15	0.02	0.28	0.02	NA	NA
Weighted mode	69	0.16	0.03	0.28	0.02	NA	NA

β: Beta estimate; CI: Confidence interval; NA: Not applicable; Q: Cochrane's Q-statistic; SNP: Single-nucleotide polymorphism.

Table 4. Causal estimates for LDL cholesterol and HDL cholesterol on ER+ breast cancer-specific survival.

Method	SNPs	HR	Lower 95% CI	Upper 95% CI	p-value	Q	Q p-value
LDL cholesterol on ER+ breast cancer-specific survival							
Inverse variance weighted	76	0.81	0.67	0.97	0.02	72	0.59
MR Egger	76	0.78	0.59	1.02	0.08	71	0.56
Weighted median	76	0.81	0.60	1.08	0.15	NA	NA
Weighted mode	76	0.79	0.61	1.02	0.08	NA	NA
HDL cholesterol on ER+ breast cancer-specific survival							
Inverse variance weighted	84	0.96	0.76	1.21	0.73	56	0.99
MR Egger	84	0.76	0.49	1.17	0.21	54	0.99
Weighted median	84	0.93	0.66	1.32	0.68	NA	NA
Weighted mode	84	0.86	0.59	1.26	0.44	NA	NA

CI: Confidence interval; HDL: High-density lipoprotein; HR: Hazards ratio; LDL: Low-density lipoprotein; NA: Not applicable; Q: Cochrane's Q-statistic; SNP: Single-nucleotide polymorphism.

Cochrane's Q-statistic demonstrated no evidence for pleiotropy. There was consistency in the direction and magnitude of the estimates across the MR sensitivity estimators.

The MR analysis of HDL cholesterol on ER+ breast cancer-specific survival was null: IVW estimate (HR) per 1-SD increase in HDL cholesterol = 0.96; 95% CI: 0.76, 1.21; p = 0.73) (Table 4). The MR-intercept (intercept = 1.01; 95% CI: 0.99, 1.03; p = 0.21) and test for heterogeneity with Cochrane's Q-statistic demonstrated no evidence for pleiotropy. There was consistency in the direction but not the magnitude of estimates across the MR sensitivity estimators.

Discussion

MR, as an analytic approach to causality, comes with caveats. The findings here must be interpreted in light of what follows: The IVW results point to a potentially protective role of higher circulating sphingomyelins against death for those with ER+ breast cancer. Supporting this, the MR-Egger sensitivity estimator confirmed the protective effect, and its intercept test provided no evidence for directional pleiotropy. However, the validity of MR-Egger regression relies also on what is known as the 'InSIDE assumption' – in this case, that no association exists between the strength of sphingomyelin-associated SNPs and the strength of bias due to horizontal pleiotropy [31]. One way to consider whether a violation to the InSIDE assumption has occurred is knowing whether a pleiotropic pathway is a potential confounder.

The PhenoScanner look-up revealed the sphingomyelin instruments were pleiotropically associated with various other lipids, including LDL cholesterol and HDL cholesterol. The subsequent statistical sensitivity analyses examining the relationships between LDL cholesterol on sphingomyelins and LDL cholesterol on and ER+ breast cancer-specific survival demonstrate that LDL cholesterol is a candidate confounder: it is genetically associated with both sphingomyelin levels and associated with ER+ breast cancer-specific survival. This is the most substantial limitation of the present study: the inability to distinguish between the effects of LDL cholesterol and sphin-

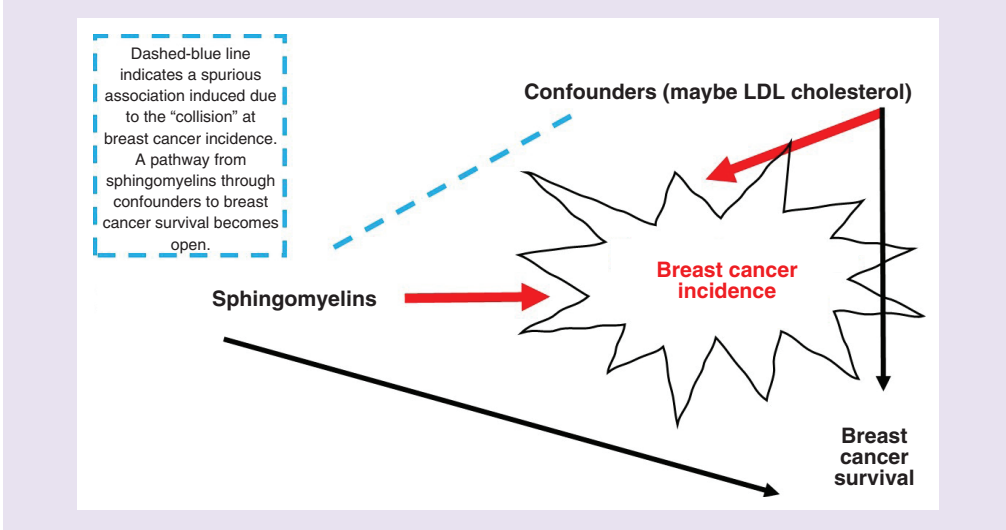


Figure 4. Hypothetical collider bias induced by breast cancer incidence being a joint cause of sphingomyelins and a confounder of breast cancer incidence and breast cancer survival.
LDL: Low-density lipoprotein.

gomyelins on ER+ breast cancer-specific survival. Unfortunately, multivariable MR, which permits adjustment, possibly untangling the relationships, is not suitable. Since the traits overlap, adjusting would render the estimates meaningless: multivariable MR would abrogate the underlying relationships by, in essence, adjusting for the same trait [32,33].

While the inability to tease out the direct effects of sphingomyelins and LDL cholesterol precludes definitive statements about either sphingomyelins or LDL protecting against death in those with ER+ breast cancer, the findings are biologically plausible. Purwaha *et al.* recently found that higher levels of sphingomyelins in tumor samples associated with better disease-free survival in triple-negative breast cancer patients [34]. They hypothesize that ceramide synthesis, through the hydrolysis of sphingomyelin, may promote progression. Similarly, Knapp *et al.* (1991) observed breast cancer patients with boney metastases had lower circulating LDL cholesterol levels than controls, implying that higher LDL cholesterol levels may be protective against progression [35].

It is also worth bearing in mind that studies involving survival are generally susceptible to a special kind of confounding known as 'collider bias' [36,37]. When studying a selected population (i.e., case-only individuals), disease incidence can be a 'collider'. The metaphor comes from visualizing the problem with a directed acyclic graph (DAG), which we will now do by drawing one in our minds. Suppose that sphingomyelins are a risk factor for breast cancer incidence (imagine a bold-red arrow going from sphingomyelins into breast cancer incidence). Also note that breast cancer incidence and breast cancer survival likely share a set of confounders (imagine another bold-red arrow, this time going from the joint confounders into breast cancer incidence). Given the two arrows both headed into breast cancer incidence, there is a potential problem ('collision') in the causal pathway at breast cancer incidence. This can induce a spurious association between the sphingomyelins and breast cancer survival through the shared confounders (Figure 4).

Fortunately, there is a way to check the plausibility of this crash. While we assumed that sphingomyelins are a risk factor for breast cancer incidence to picture collider bias, if sphingomyelins are not associated with breast cancer incidence, then the potential collision at breast cancer incidence does not happen. Recently, the relationship between sphingomyelins and breast cancer incidence was investigated and found to be null [38]. As such, the risk for confounding due to collider bias is reduced. This does not eliminate other sources of confounding previously mentioned.

Both sphingomyelins and LDL cholesterol deserve future attention in the context of breast cancer progression and survival. Replication of the present findings, when more summary data for mortality are available, is needed.

Moreover, bias from pleiotropy cannot be fully ruled out in any MR study. Solid causal inference is made through triangulating multiple lines of evidence – building a case for causality that gathers evidence from studies that do not suffer from the same possible biases.

Future perspective

Therefore, in addition to replicating with MR, future investigations should triangulate the present study's results by exploring the role of these lipids on breast cancer-specific survival in designs other than MR. Should the findings for the protective effect of sphingomyelins on ER+ breast cancer-specific survival replicate and remain robust when examined with other methods, looking into ways to use sphingomyelins therapeutically to manage ER+ breast cancer becomes a promising avenue for future research.

Summary points

- If a causal relationship exists between circulating sphingomyelins and breast-cancer specific survival, circulating sphingomyelins could be studied as a therapeutic target in the management of breast cancer.
- Mendelian randomization (MR) is a quasi-randomization, analytic technique that can be used to study causal relationships in epidemiologic data.
- MR exploits Mendel's laws of inheritance and that genotypes are set at conception, temporally preceding observational factors of interest.
- Due to the reliance on Mendel's laws of inheritance and genotype assignment at conception, Mendelian randomization avoids most sources of confounding and reverse causation, and, for that reason, is referred to as a 'causal' method.
- MR uses genetic variants that proxy for a trait of interest in statistical models in lieu of the trait itself.
- This MR study, which used genetic variants proxying for sphingomyelins instead of circulating sphingomyelins in its model, found evidence that circulating sphingomyelins are protective for ER+ breast-cancer specific survival but null ER− breast-cancer specific survival.
- Sensitivity analyses revealed that low-density lipoprotein cholesterol may contribute to the findings and/or be a confounder.
- Future replication studies and triangulation with methods other than Mendelian randomization are needed to confirm or refute the present observations.

Supplementary data

To view the supplementary data that accompany this paper please visit the journa website at: www.futuremedicine.com/doi/suppl/10.2217/bmt-2020-0002

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J. Clin.* 68(1), 7–30 (2018).
2. Nagahashi M, Tsuchida J, Moro K *et al.* High levels of sphingolipids in human breast cancer. *J. Surg. Res.* 204(2), 435–444 (2016).
3. Sobhani N, Roviello G, Corona SP *et al.* The prognostic value of PI3K mutational status in breast cancer: a meta-analysis. *J. Cell Biochem.* 119(6), 4287–4292 (2018).
4. Oskouian B, Saba JD. Cancer treatment strategies targeting sphingolipid metabolism. *Adv. Exp. Med. Biol.* 688, 185–205 (2010).
5. Stearns V, Park BH. PIK3CA mutations in hormone receptor-positive breast cancers: PIKing biomarkers to inform adjuvant endocrine therapy decisions. *JAMA Oncol.* 4(10), 1330–1332 (2018).
6. Ryland LK, Fox TE, Liu X, Loughran TP, Kester M. Dysregulation of sphingolipid metabolism in cancer. *Cancer Biol. Ther.* 11(2), 138–149 (2011).
7. Schiffmann S, Sandner J, Birod K *et al.* Ceramide synthases and ceramide levels are increased in breast cancer tissue. *Carcinogenesis* 30(5), 745–752 (2009).

8. Moro K, Kawaguchi T, Tsuchida J *et al.* Ceramide species are elevated in human breast cancer and are associated with less aggressiveness. *Oncotarget* 9(28), 19874–19890 (2018).
9. Ogretmen B, Hannun YA. Biologically active sphingolipids in cancer pathogenesis and treatment. *Nat. Rev. Cancer* 4(8), 604–616 (2004).
10. Gray R, Wheatley K. How to avoid bias when comparing bone marrow transplantation with chemotherapy. *Bone Marrow Transpl.* 7(Suppl 3), 9–12 (1991).
11. Adams D. The gene-based hack that is revolutionizing epidemiology. *Nature News Feature* (2019). www.nature.com/articles/d41586-019-03754-3
12. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.* 23(R1), R89–R98 (2014).

•• **This is a foundational paper for understanding Mendelian randomization (MR).**

13. Smith GD, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* 32(1), 1–22 (2003).

•• **This is a foundational paper for understanding MR.**

14. Kettunen J, Demirkan A, Wurtz P *et al.* Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nat. Commun.* 7, 1–9 (2016).
15. MRBASE. 2-sample Mendelian Randomisation (2018). www.mrbase.org/
16. Hemani G, Zheng J, Elsworth B *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 7, 1–29 (2018).
17. Guo Q, Schmidt MK, Kraft P *et al.* Identification of novel genetic markers of breast cancer survival. *J. Natl Cancer Inst.* 107(5), 1–9 (2015).
18. Brion MA, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int. J. Epidemiol.* 42, 1497–1501 (2013).
19. mRnd: Power calculations for Mendelian Randomization. https://cnsgenomics.shinyapps.io/mRnd/
20. Yarmolinsky J, Relton CL, Lophatananon A *et al.* Appraising the role of previously reported risk factors in epithelial ovarian cancer risk: a Mendelian randomization analysis. *PLoS Med.* 16(8), e1002893 (2019).

•• **The supplement to this article provides nuanced descriptions of the MR methods and assumptions.**

21. Del Greco MF, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat. Med.* 34(21), 2926–2940 (2015).
22. Staley JR, Blackshaw J, Kamat MA *et al.* PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* 32(20), 3207–3209 (2016).
23. Kamat MA, Blackshaw JA, Young R *et al.* PhenoScanner V2: an expanded tool for searching human genotype–phenotype associations. *Bioinformatics* 35(22), 4851–4853 (2019).
24. Hanamatsu H, Ohnishi S, Sakai S *et al.* Altered levels of serum sphingomyelin and ceramide containing distinct acyl chains in young obese adults. *Nutr. Diabetes* 4(10), e141–e147 (2014).
25. Deevska GM, Sunkara M, Morris AJ, Nikolova-Karakashian MN. Characterization of secretory sphingomyelinase activity, lipoprotein sphingolipid content and LDL aggregation in ldlr-/- mice fed on a high-fat diet. *Biosci. Rep.* 32(5), 479–490 (2012).
26. Martínez-Beamonte R, Lou-Bonafonte JM, Martínez-Gracia MV, Osada J. Sphingomyelin in high-density lipoproteins: structural role and biological function. *In. J. Mol. Sci.* 14(4), 7716–7741 (2013).
27. Hammad SM, Pierce JS, Soodavar F *et al.* Blood sphingolipidomics in healthy humans: impact of sample collection methodology. *J. Lipid Res.* 51(10), 3074–3087 (2010).
28. Nilsson Å, Duan RD. Absorption and lipoprotein transport of sphingomyelin. *J. Lipid Res.* 47(1), 154–171 (2006).
29. Bowden J, Spiller W, M FDG *et al.* Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the radial plot and radial regression. *Int. J. Epidemiol.* 47(4), 1264–1278 (2018).
30. Willer CJ, Schmidt EM, Sengupta S *et al.* Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* 45(11), 1274–1283 (2013).
31. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur. J. Epidemiol.* 32(5), 377–389 (2017).
32. Holmes M V, Davey Smith G. Challenges in interpreting multivariable Mendelian randomization: might "good cholesterol" be good after all? *Am. J. Kidney Dis.* 71(2), 149–153 (2018).
33. Holmes MV, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat. Rev. Cardiol.* 14(10), 577–599 (2017).

34. Purwaha P, Gu F, Piyarathna DWB *et al.* Unbiased lipidomic profiling of triple-negative breast cancer tissues reveals the association of sphingomyelin levels with patient disease-free survival. *Metabolites* 8(3), 1–14 (2018).

35. Knapp ML, Al-Sheibani S, Riches PG. Alterations of serum lipids in breast cancer: effects of disease activity, treatment, and hormonal factors. *Clin. Chem.* 37(12), 2093–2101 (1991).

36. Munafò MR, Tilling K, Taylor AE, Evans DM, Smith GD. Collider scope: when selection bias can substantially influence observed associations. *Int. J. Epidemiol.* 47(1), 226–235 (2018).

•• **Provides a detailed explanation of collider bias in MR studies of progression.**

37. Paternoster L, Tilling K, Davey Smith G. Genetic epidemiology and Mendelian randomization for informing disease therapeutics: conceptual and methodological challenges. *PLoS Genet.* 13(10), 1–9 (2017).

•• **Provides a detailed explanation of collider bias in MR studies of progression.**

38. Adams CD. Null effect of circulating sphingomyelins on risk for breast cancer: a Mendelian randomization report using Breast Cancer Association Consortium (BCAC) data. *medRxiv* 31(0), 1–26 (2019).

• **Provides the null results for sphingomyelins on breast cancer initiation, which is helpful for assessing the likelihood of collider bias in the present study.**