compared with miR-33$^{+/+}$*Apoe*$^{-/-}$ mice (P=0.0066; Figure 3A and 3B). The αSMA-positive area was also significantly reduced in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice (P=0.025 Figure 3C and 3D). However, no difference was noted in the collagen area (Figure 3E and 3F). There was a significant reduction in lesional apoptosis, as measured by anti-ssDNA staining (P=0.0493; Figure 3G and 3H). These results indicated that a deficiency of miR-33 decreased atherosclerotic plaque size and lipid content and reduced the accumulation of macrophages and T cells in atherosclerotic plaques.

MiR-33 Deficiency Increased HDL-C

We previously reported that miR-33$^{-/-}$ mice showed 22% to 39% higher serum HDL-C levels than wild-type mice.[22] We measured HDL-C levels in the serum of miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice at the time of euthanization by the standard method. HDL-C was significantly elevated in miR-33$^{-/-}$ *Apoe*$^{-/-}$ mice compared with miR-33$^{+/+}$*Apoe*$^{-/-}$ mice of both sexes (Figure 4A). We further classified and quantified serum lipoproteins by high-performance liquid chromatography (HPLC). Representative results of the HPLC elution profile of serum of both sexes are shown in Figure 4B and 4C, and lipid profiles are summarized in Table 1. These results show that only the HDL-C level differed between the serum of miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Serum apoA-I levels are shown in Figure 4D. To assess the cholesterol efflux capacity of the serum, cholesterol efflux via apolipoprotein B (apoB)–depleted serum was measured using ^{3}H-cholesterol-labeled J774 mouse macrophages. The mean values of serum HDL-C used in this experiment were 17.8±1.4 versus 19.5±1.8 mg/dL in males and 10.2±0.5 versus 17.3±1.7 mg/dL in females (miR-33$^{+/+}$ *Apoe*$^{-/-}$ versus miR-33$^{-/-}$*Apoe*$^{-/-}$ mice, n=6 for each group). ApoB-depleted serum from miR-33$^{-/-}$ *Apoe*$^{-/-}$ mice significantly promoted cholesterol efflux in J774 macrophages (Figure 4E). These results indicated that deficiency of miR-33 elevated serum cholesterol efflux capacity, possibly through the elevation of HDL-C levels.

Peritoneal Macrophages From MiR-33$^{-/-}$*Apoe*$^{-/-}$ Mice Showed Improved Cholesterol Efflux

To characterize the function of macrophages in cholesterol efflux, thioglycollate-elicited peritoneal macrophages (PEMs) were isolated from mice. Previously, we and others have shown that miR-33 targeted the 3′ untranslated region (UTR) of *Abca1* and *Abcg1*.[16–18,22] mRNA expression of ABCA1 and protein expression of ABCA1 and ABCG1 were significantly increased in

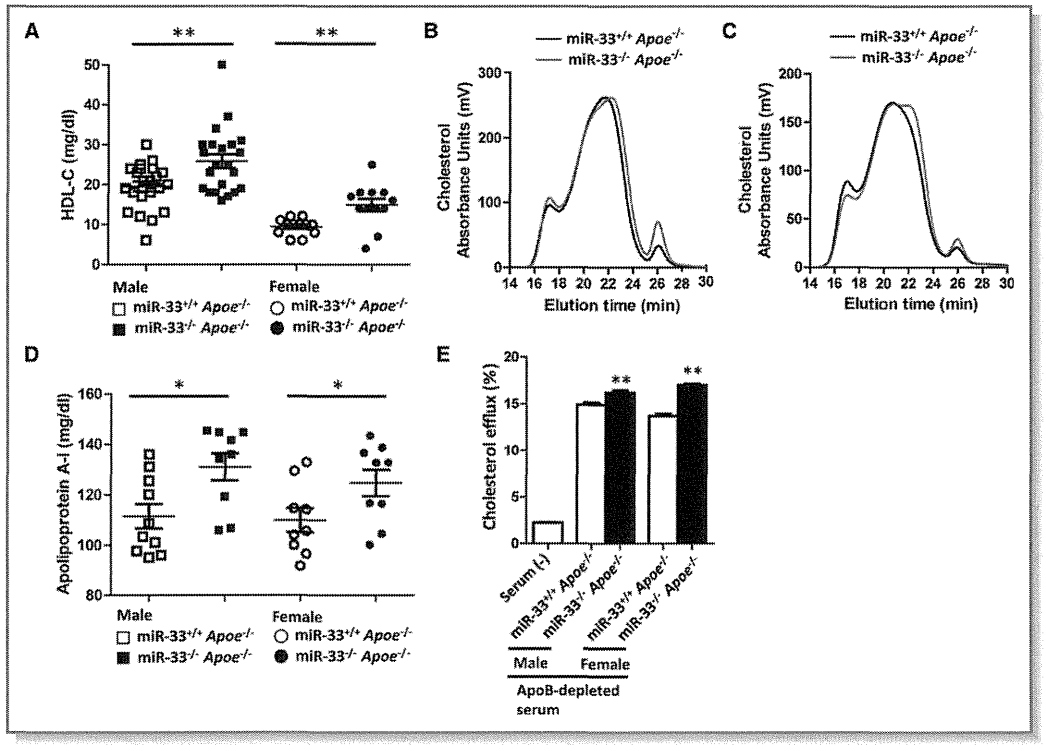

Figure 4. miR-33 deficiency increased HDL-C. A, Serum HDL-C levels determined by standard methods in miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$ *Apoe*$^{-/-}$ male and female mice. Values are mean±SE (male, n=22; female, n=11 each); **P<0.01. B, Representative HPLC analysis of serum cholesterol from male miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. C, Representative HPLC analysis of serum cholesterol from female miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. D, Serum apoA-I levels in miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice; *P<0.05. E, Cholesterol efflux via apoB-depleted serum from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice using ^{3}H-cholesterol-labeled J774 mouse macrophages. Values are mean±SE (n=6 each); **P<0.01. HDL-C indicates high-density lipoprotein cholesterol; HPLC, high-performance liquid chromatography.

Table 1. Serum Lipid Profiling of MiR-33$^{+/+}$*Apoe*$^{-/-}$ and MiR-33$^{-/-}$*Apoe*$^{-/-}$ Mice by HPLC

	Lipoprotein		Male		Female	
	Major (Fraction No.) Diameter	Subclass (Fraction No.)	miR-33$^{+/+}$ *Apoe*$^{-/-}$	miR-33$^{-/-}$ *Apoe*$^{-/-}$	miR-33$^{+/+}$ *Apoe*$^{-/-}$	miR-33$^{-/-}$ *Apoe*$^{-/-}$
TC, mg/dL			705.3±72.1	768.4±45.2	598.9±88.8	574.2±70.0
	CM (1 to 2), >80 nm		129.2±21.6	157.9±19.8	124.6±20.9	138.6±28.11
	VLDL (3 to 7), 30 to 80 nm		419.5±44.1	450.9±30.0	353.6±51.7	322.1±36.4
		Large VLDL (3 to 5)	279.3±32.2	316.4±24.9	245.1±35.1	230.2±29.5
		Medium VLDL (6)	100.5±9.8	96.77±5.67	79.04±12.17	66.37±6.15
		Small VLDL (7)	39.74±3.31	37.75±1.47	29.47±4.87	25.6±2.12
	LDL (8 to 13), 16 to 30 nm		133.8±14.9	126.4±5.23	103.0±16.2	90.81±6.72
		Large LDL (8)	47.47±4.40	44.05±1.53	36.58±6.03	31.60±2.61
		Medium LDL (9)	39.93±4.56	36.35±1.71	31.99±5.26	27.53±2.11
		Small LDL (10)	25.63±3.42	24.72±1.49	19.78±3.05	17.36±1.23
		Very small LDL (11 to 13)	20.75±2.75	21.25±1.19	14.65±1.89	14.33±1.27
	HDL (14 to 20), 8 to 16 nm		22.79±1.28	33.21±2.51**	17.62±0.73	22.71±2.52*
		Very large HDL (14 to 15)	3.04±0.39	3.66±0.17	3.57±0.42	3.89±0.56
		Large HDL (16)	5.88±0.63	9.39±0.87**	4.00±0.28	5.29±0.47*
		Medium HDL (17)	8.04±0.71	11.88±1.21*	4.49±0.36	6.79±0.26**
		Small HDL (18)	2.61±0.12	3.73±0.30**	2.16±0.13	2.87±0.36
		Very small HDL (19 to 20)	3.23±0.25	4.55±0.27**	3.40±0.45	3.87±0.61
TG, mg/dL			39.94±8.29	48.32±10.14	17.60±1.44	24.75±5.19

Values are mean±SEM. After a 4-hour fast, blood was obtained from mice fed a WTD for 16 weeks (male, n=6 each; female, n=5 each). The serum was analyzed by HPLC, as described in Methods. TC indicates total cholesterol; TG, triglyceride; CM, chylomicrons; VLDL, very low-density lipoprotein; LDL, low-density lipoprotein; HDL, high-density lipoprotein.
*P<0.05, **P<0.01 compared with miR-33$^{+/+}$*Apoe*$^{-/-}$ mice.

PEMs from miR-33$^{-/-}$*Apoe*$^{-/-}$ mice compared with PEMs from miR-33$^{+/+}$*Apoe*$^{-/-}$ mice (Figure 5A and 5B). To test the hypothesis that enhanced cholesterol efflux in macrophages contributed to the reduction in the development of atherosclerosis in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice, cholesterol efflux to apoA-I and HDL-C was measured using ^{3}H-cholesterol-labeled acetylated low-density lipoprotein (acLDL). Both apoA-I- and HDL-C-mediated cholesterol efflux were significantly elevated in macrophages from miR-33$^{-/-}$*Apoe*$^{-/-}$ compared with macrophages from miR-33$^{+/+}$*Apoe*$^{-/-}$ mice in a dose-dependent manner (Figure 5C and 5D). These results indicate that miR-33 deficiency improved macrophage cholesterol efflux by increasing the expressions of macrophage ABCA1 and ABCG1.

MiR-33 Deficiency Affected Circulating Monocytes and Lesional Macrophages

Our findings thus far indicate that miR-33 deficiency reduced the accumulation of inflammatory cells in atherosclerotic plaque. We then examined whether miR-33 deficiency influenced the monocyte count or subset frequency in peripheral blood. The total leukocyte count in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice was significantly less than that in miR-33$^{+/+}$*Apoe*$^{-/-}$ mice (Figure 6A). Blood monocyte subsets were discriminated by flow cytometry on the basis of their expression of CD115 and Ly6C.[36] The frequency of proinflammatory Ly6C^{high} monocytes in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice was significantly higher than that in miR-33$^{+/+}$*Apoe*$^{-/-}$ mice (Figure 6B through 6E). We further measured iNOS, interleukin (IL)–6, and IL-10 expression in atherosclerotic plaque by immunohistochemistry. As shown in Figure 7A and 7B, the iNOS-positive area in atherosclerotic plaque in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice was significantly less than that in miR-33$^{+/+}$*Apoe*$^{-/-}$ mice (P=0.0057). On the other hand, no differences were observed in IL-6 and IL-10 immunostaining (Figure 7C through 7F). We also characterized miR-33$^{-/-}$*Apoe*$^{-/-}$ macrophages by analyzing the expression of classically activated or proinflammatory (M1) and alternatively activated or anti-inflammatory (M2) macrophage markers using mouse PEMs. A previous article indicated that the inhibition of miR-33 by antisense oligonucleotide enhanced M2 marker expression in macrophages.[19] However, our experiments indicated that the mRNA levels of both M2 markers such as *IL-10* and *Chi3l3* mRNA and an M1 marker such as *IL-6* in PEMs from miR-33$^{-/-}$*Apoe*$^{-/-}$ mice were significantly elevated compared with those from miR-33$^{+/+}$*Apoe*$^{-/-}$ mice (Figure 7G). Overall, these results

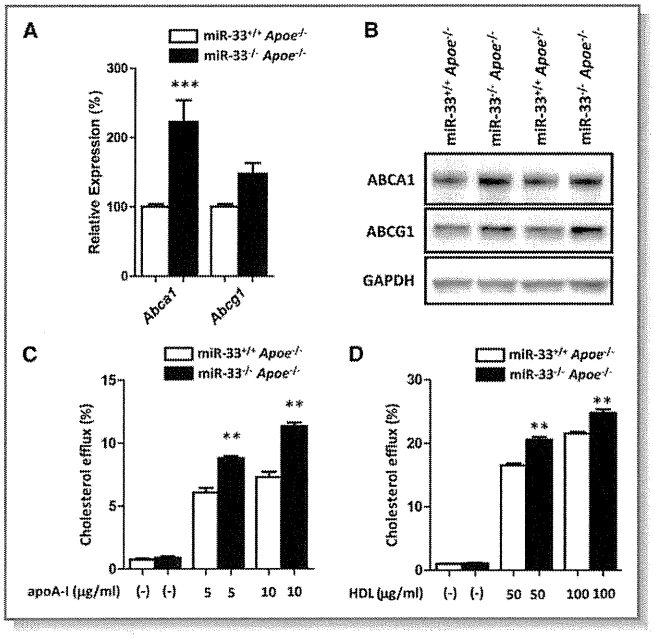

Figure 5. miR-33 deficiency improved cholesterol efflux in macrophages. A, Quantitative real-time PCR analysis of *Abca1* and *Abcg1* in macrophages from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Values from miR-33$^{+/+}$*Apoe*$^{-/-}$ were set at 100%. Values are mean±SE (n=7 each); ***P<0.001. B, Western blotting analysis of ABCA1 and ABCG1 in thioglycollate-elicited peritoneal macrophages from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. GAPDH was used as a loading control. C, Cholesterol efflux from thioglycollate-elicited peritoneal macrophages in the presence or absence of apoA-I (5 or 10 μg/mL). Values are mean±SE (n=6 each); **P<0.01. D, Cholesterol efflux from thioglycollate-elicited peritoneal macrophages in the presence or absence of HDL-C (50 or 100 μg/mL). Values are mean±SE (n=6 each); **P<0.01.

demonstrate that a loss of miR-33 may have affected multiple pathways in both pro- and anti-inflammatory processes. Moreover, we analyzed adhesion molecule expression by immunostaining. As shown in Figure 8A and 8B, the VCAM-1-positive area in atherosclerotic plaque in miR-33$^{-/-}$*Apoe*$^{-/-}$ was significantly less than that in miR-33$^{+/+}$*Apoe*$^{-/-}$ mice (P=0.0008). The ICAM-1-positive area in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice tended to be less than that in miR-33$^{+/+}$*Apoe*$^{-/-}$ mice (P=0.127), shown in Figure 8C and 8D.

MiR-33 Target Genes Were Altered in the Livers of MiR-33$^{-/-}$ Mice on an *Apoe*$^{-/-}$ Background

It has already been shown that miR-33 targets several genes that affect cholesterol and fatty acid synthesis. We measured the mRNA and protein levels of ABCA1, CROT, and CPT1a in the livers of these mice. As shown in Figure 9A through 9C, mRNA of CROT and protein levels of ABCA1 and CROT in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice were significantly elevated compared with those in miR-33$^{+/+}$*Apoe*$^{-/-}$ mice. No differences were observed in CPT1a and AMPKα expressions. Next, we measured the lipid content in the liver. However, there was no difference in total cholesterol, free cholesterol, cholesterol ester, or triglyceride levels in the livers of miR-33$^{+/+}$*Apoe*$^{-/-}$ mice and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice (Figure 9D). No apparent changes in histology were observed in the livers of these mice, as shown by HE staining (Figure 9E). Moreover, we measured the level of RIP140 (NRIP1), which has been shown to be one of the targets of miR-33 in macrophages of these mice.[37] Protein level of RIP140 in miR-33$^{-/-}$*Apoe*$^{-/-}$ macrophages was significantly increased compared with that in miR-33$^{+/+}$*Apoe*$^{-/-}$ macrophages (Figure 9F through 9H), which may be one of the reasons why the expression of inflammatory cytokine such as IL-6 in PEMs of miR-33$^{-/-}$*Apoe*$^{-/-}$ mice was increased compared with that in miR-33$^{+/+}$*Apoe*$^{-/-}$ mice.[38]

Loss of MiR-33 in Blood Cells Did Not Alter Serum HDL-C Levels

The results of these experiments show that both the rise in HDL-C level and the improvement in macrophage cholesterol efflux may have contributed to the reduction in atherosclerotic plaques. To elucidate the contribution of miR-33 in macrophages to the development of atherosclerosis in vivo, we used bone marrow transplantation (BMT) to generate *Apoe*$^{-/-}$ mice selectively deficient in leukocyte miR-33. Male mice with miR-33$^{+/+}$*Apoe*$^{-/-}$ or miR33$^{-/-}$*Apoe*$^{-/-}$ genotypes (8 weeks old) were used as bone marrow (BM) donors. BM recipients were female miR-33$^{+/+}$*Apoe*$^{-/-}$ mice (8 weeks old). Thus, all the mice used for BMT had an *Apoe*$^{-/-}$ background. After BMT, mice were fed NC for 4 weeks and then switched to a WTD for 12 weeks. At age 24 weeks, mice were euthanized and analyzed (Figure 10A). Successful hematopoietic reconstitution after BMT was confirmed by PCR amplification of the whole-blood genome 4 weeks after BMT and by PCR amplification of the BM and tail genomes at the time of euthanization. MiR-33 was barely detectable in BM cells from miR-33$^{-/-}$*Apoe*$^{-/-}$ BM recipients by quantitative PCR analysis for miR-33 (data not shown). The serum lipid profile of recipient mice is shown in Table 2. Serum HDL-C levels of miR-33$^{+/+}$*Apoe*$^{-/-}$ BM recipients were the same as those in miR-33$^{-/-}$*Apoe*$^{-/-}$ BM recipients, which were similar to the levels in miR-33$^{+/+}$*Apoe*$^{-/-}$ mice in Figure 4A. These results indicated that miR-33 expression in macrophages did not contribute to serum HDL-C levels, which is consistent with the results that the liver and intestine are the major sources of HDL-C.[39,40]

MiR-33$^{+/+}$*Apoe*$^{-/-}$ Mice Transplanted With MiR-33$^{-/-}$*Apoe*$^{-/-}$ BM Had Reduced Lipid Accumulation in Atherosclerotic Plaque

After mice were fed a WTD for 12 weeks, atherosclerotic lesions in the proximal aortas were measured. Atherosclerotic

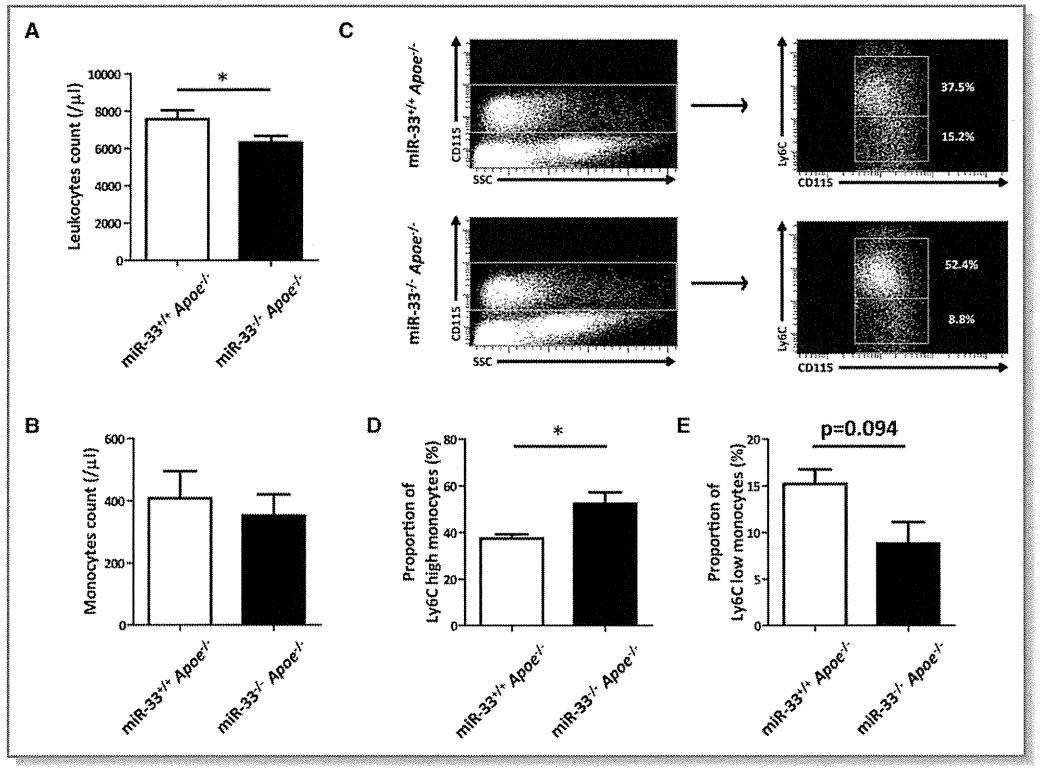

Figure 6. miR-33$^{-/-}$*Apoe*$^{-/-}$ mice had reduced leukocyte numbers in peripheral blood and a greater Ly-6C^{high} monocyte subset compared with miR-33$^{+/+}$*Apoe*$^{-/-}$ mice. A, Leukocyte count in peripheral blood in miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Values are mean±SE (n=15 to 16 each); *P<0.05. B, Numbers of monocyte in peripheral blood in miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Values are mean±SE (n=9 each). C, Scheme for gating of monocytes using an anti-CD115 antibody, and representative dot plots showing the quantification of Ly-6C^{high} and Ly-6C^{low} monocyte subsets in miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. D, Proportion of the Ly-6C^{high} monocyte subset to total monocytes in miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Values are mean±SE (n=9 each.); *P<0.05. E, Proportion of the Ly-6C^{low} monocyte subset to total monocytes in miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Values are mean±SE (n=9 each).

plaque formation in mice transplanted with miR-33$^{-/-}$*Apoe*$^{-/-}$ BM tended to be reduced compared with that in mice transplanted with miR-33$^{+/+}$*Apoe*$^{-/-}$ BM, but this difference was not statistically significant (P=0.12, 0.53±0.028 versus 0.46±0.033 mm^2; Figure 10B and 10C). On the other hand, the quantification of lipid accumulation by oil red O staining showed a significant decrease in mice transplanted with miR-33$^{-/-}$*Apoe*$^{-/-}$ BM compared with mice with miR-33$^{+/+}$ *Apoe*$^{-/-}$ BM (P=0.020, 3.0±0.4% versus 1.9±0.2%; Figure 10D and 10E). The CD68-positive area is shown in Figure 10F and 10G (P=0.09, 19.5±2.6% versus 26.6±2.6%). These results showed that loss of miR-33 in blood cells reduced the lipid content of atherosclerotic plaque.

Free Cholesterol–Induced Apoptosis Was Reduced in PEMs From MiR-33$^{-/-}$*Apoe*$^{-/-}$ Mice Compared With MiR-33$^{+/+}$*Apoe*$^{-/-}$ Mice

Because mice transplanted with miR-33$^{-/-}$*Apoe*$^{-/-}$ BM showed reduced lipid accumulation in atherosclerotic plaque compared with mice transplanted with miR-33$^{+/+}$*Apoe*$^{-/-}$ BM, we analyzed free cholesterol (FC)–induced apoptosis in PEMs by treating macrophages with acLDL and acyl-CoA: cholesterol acyl-transferase (ACAT) inhibitor. Figure 11A shows the results of annexin V staining of PEMs with and without acLDL plus ACAT inhibitor treatment. Annexin V–positive cells were significantly increased in PEMs from miR-33$^{+/+}$*Apoe*$^{-/-}$ mice compared with those from miR-33$^{-/-}$ *Apoe*$^{-/-}$ mice after FC loading (Figure 11B). Cleaved caspase-3 was also increased in PEMs from miR-33$^{+/+}$*Apoe*$^{-/-}$ mice compared with those from miR-33$^{-/-}$*Apoe*$^{-/-}$ mice after FC loading (Figure 11C).

MiR-33$^{-/-}$*Apoe*$^{-/-}$ Mice Transplanted With BM of MiR-33$^{+/+}$*Apoe*$^{-/-}$ or MiR-33$^{-/-}$*Apoe*$^{-/-}$ Showed Only a Slight Increase in HDL-C Levels Compared With MiR-33$^{+/+}$*Apoe*$^{-/-}$ Mice Transplanted With BM of MiR-33$^{+/+}$*Apoe*$^{-/-}$ or MiR-33$^{-/-}$*Apoe*$^{-/-}$

Previous experiments indicated that loss of miR-33 in blood cells reduced lipid accumulation in atherosclerotic plaque.

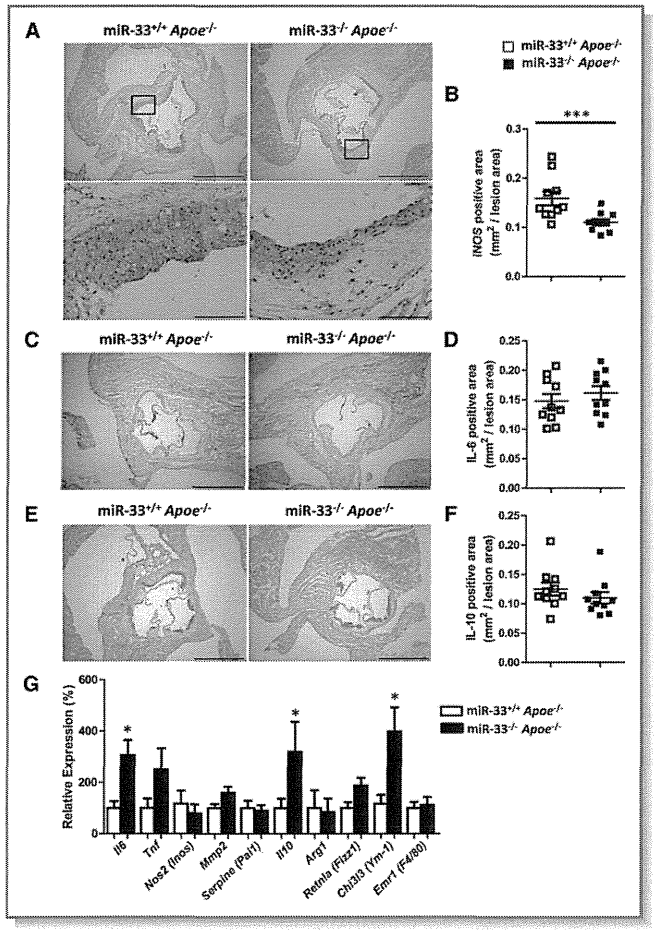

Figure 7. miR-33 deficiency reduced the iNOS-positive areas in atherosclerotic plaque and induced coordinated M1 and M2 marker expression in PEMs. A, Representative microscopic images of immunohistochemical staining for iNOS in male mice. Scale bars: 1 mm (upper), 100 μm (lower). B, Quantification of the iNOS-positive area in cross-sections of proximal aorta in male mice. Values are mean±SE (n=10 each); ***P<0.001. C, Representative microscopic images of immunohistochemical staining for IL-6 in male mice. Scale bar: 1 mm. D, Quantification of the IL-6-positive area in cross-sections of proximal aorta in male mice. Values are mean±SE (n=10 each). E, Representative microscopic images of immunohistochemical staining for IL-10 in male mice. Scale bar: 1 mm. F, Quantification of the IL-10-positive area in cross-sections of proximal aorta in male mice. Values are mean±SE (n=10 each). G, Quantitative real-time PCR analysis of proinflammatory (M1) and anti-inflammatory (M2) markers in residual PEMs from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Values from miR-33$^{+/+}$*Apoe*$^{-/-}$ mice were set at 100%. Values are mean±SE (n=7 each); *P<0.05.

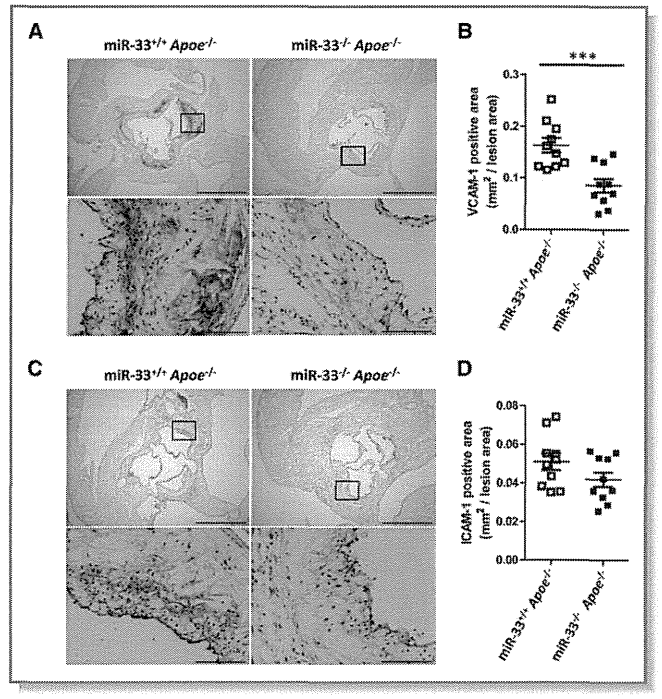

Figure 8. miR-33 deficiency reduced the VCAM-1-positive area in atherosclerotic plaque. A, Representative microscopic images of immunohistochemical staining for VCAM-1 in male mice. Scale bars: 1 mm (upper), 100 μm (lower). B, Quantification of the VCAM-1-positive area in cross-sections of proximal aorta in male mice. Values are mean±SE (n=10 each); ***P<0.001. C, Representative microscopic images of immunohistochemical staining for ICAM-1 in male mice. Scale bars: 1 mm (upper), 100 μm (lower). D, Quantification of the ICAM-1-positive area in cross-sections of proximal aorta in male mice. Values are mean±SE (n=10); P=0.127.

Next, to determine the contribution of miR-33 deficiency to atherosclerosis in BM recipients, we transferred BM of miR-33$^{+/+}$*Apoe*$^{-/-}$ or miR-33$^{-/-}$*Apoe*$^{-/-}$ mice to miR-33$^{-/-}$*Apoe*$^{-/-}$ mice in the same way as in the previous BMT experiments (Figure 12A). However, the HDL-C levels in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice transplanted with BM of miR-33$^{+/+}$*Apoe*$^{-/-}$ or miR-33$^{-/-}$*Apoe*$^{-/-}$ (Table 3) showed only a slight increase compared with miR-33$^{+/+}$*Apoe*$^{-/-}$ mice transplanted with BM of miR-33$^{+/+}$*Apoe*$^{-/-}$ or miR-33$^{-/-}$*Apoe*$^{-/-}$ (Table 2). Therefore, it was impossible to observe an effect of HDL-C elevation caused by the loss of miR-33 on atherosclerosis in recipients that had the same type of blood cells (Figures 10 and 12).

MiR-33$^{-/-}$*Apoe*$^{-/-}$ Mice Transplanted With MiR-33$^{-/-}$*Apoe*$^{-/-}$ BM Had Reduced Lipid Accumulation in Atherosclerotic Plaque

We also observed the effect of loss of miR-33 in BMT experiments in miR-33$^{-/-}$*Apoe*$^{-/-}$ recipients. Atherosclerotic plaque formation in miR-33$^{-/-}$ *Apoe*$^{-/-}$mice transplanted with miR-33$^{-/-}$*Apoe*$^{-/-}$ BM was comparable with that in miR-33$^{-/-}$ *Apoe*$^{-/-}$mice transplanted with miR-33$^{+/+}$ *Apoe*$^{-/-}$ BM (0.46±0.023 versus 0.42±0.020 mm^2; Figure 12B and 12C). Although the area of lipid accumulation in miR-33$^{-/-}$*Apoe*$^{-/-}$mice transplanted with miR-33$^{-/-}$ *Apoe*$^{-/-}$ BM was significantly reduced compared with that in miR-33$^{-/-}$*Apoe*$^{-/-}$mice transplanted with miR-33$^{+/+}$

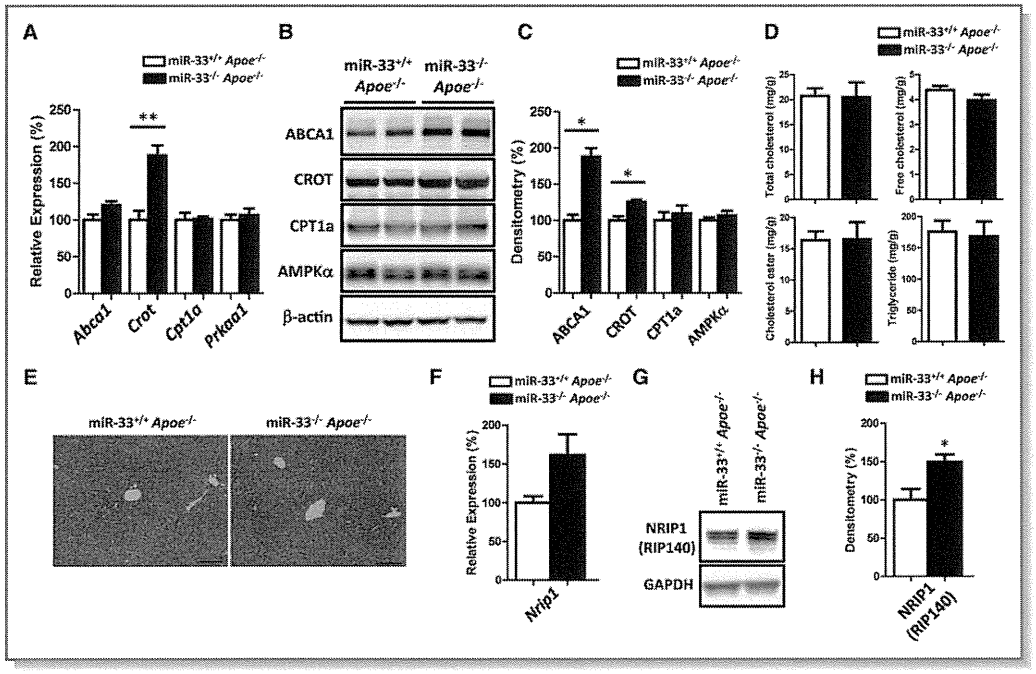

Figure 9. Expression of ABCA1 and CROT in livers and RIP140 in macrophages is elevated in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice compared with miR-33$^{+/+}$*Apoe*$^{-/-}$ mice. A, Quantitative real-time PCR analysis of *Abca1*, *Crot*, *Cpt1a*, and *Prkaa1*in livers from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Values from miR-33$^{+/+}$*Apoe*$^{-/-}$ mice were set at 100%. Values are mean±SE (n=9 to 11each); **P<0.01. B, Western analysis of ABCA1, CROT, CPT1a, and AMPKα in livers from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. β-actin was used as a loading control. C, Densitometry of ABCA1, CROT, CPT1a, and AMPKα in livers from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Values from miR-33$^{+/+}$*Apoe*$^{-/-}$ mice were set at 100%. Values are mean±SE (n=4 each); *P<0.05. D, Total cholesterol, free cholesterol, cholesterol ester, and triglyceride levels in livers of miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Values are mean±SE (n=9 to 11 each). E, HE staining of livers of miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice at age 20 weeks fed NC. Scale bar: 100 μm. F, Quantitative real-time PCR analysis of *Nrip1* (RIP140) in peritoneal macrophages from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Values from miR-33$^{+/+}$*Apoe*$^{-/-}$ mice were set at 100%. Values are mean±SE (n=7 each). G, Western analysis of NRIP1 (RIP140) in peritoneal macrophages from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. GAPDH was used as a loading control. H, Densitometry of NRIP1 (RIP140) in peritoneal macrophages from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Values from miR-33$^{+/+}$*Apoe*$^{-/-}$ mice were set at 100%. Values are mean±SE (n=4 each); *P<0.05.

Apoe$^{-/-}$ BM (P=0.027, 3.4±0.3% versus 2.2±0.3%; Figure 12D and 12E), there was no difference between these mice in the CD68-positive area (P=0.37, 35.3±3.5% versus 41.2±3.3%; Figure 12F and 12G).

Discussion

In the current study, miR-33$^{-/-}$*Apoe*$^{-/-}$ mice showed an increase in HDL-C in vivo and a decrease in atherosclerotic plaque size and lipid content compared with miR-33$^{+/+}$*Apoe*$^{-/-}$ mice. Although a previous study demonstrated that the short-term administration of anti-miR-33 oligonucleotides raised HDL-C levels and promoted the regression of atherosclerosis, this current study indicates that miR-33 deficiency contributes to the reduction of plaque size in the progression of advanced atherosclerosis. Moreover, we assessed the in vivo function of miR-33 deficiency in leukocytes by BMT from miR-33$^{+/+}$*Apoe*$^{-/-}$ or miR-33$^{-/-}$*Apoe*$^{-/-}$ mice into miR-33$^{+/+}$*Apoe*$^{-/-}$ or miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. miR-33$^{-/-}$*Apoe*$^{-/-}$ macrophages promoted the removal of intracellular lipid content compared with miR-33$^{+/+}$*Apoe*$^{-/-}$ macrophages. Together, these data demonstrate that miR-33 deficiency serves to raise HDL-C, improve cholesterol efflux in macrophages, and prevent the progression of atherosclerosis and suggest that miR-33 should be considered as a potential target to prevent the progression of atherosclerosis.

We reported previously that miR-33$^{-/-}$ in C57/BL6 mice increases HDL-C by up to 40%. MiR-33$^{-/-}$*Apoe*$^{-/-}$ mice also had a higher amount of HDL-C compared with controls. However, the static measurement of HDL-C cholesterol level has inherent limitations as a metric of the functional effects of HDL-C in vivo. Moreover, HDL-C function in *Apoe*$^{-/-}$ mice may be altered because these mice had no apoE to activate lecithin: cholesterol acyltransferase, which converts free cholesterol to cholesterol ester, thereby creating a gradient for free cholesterol efflux from cells to HDL-C. Thus, we tried to measure the function of HDL-C from miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. Recently, the functionality of HDL-C was analyzed by the quantification of efflux capacity from blood samples of humans.[25] Therefore, we

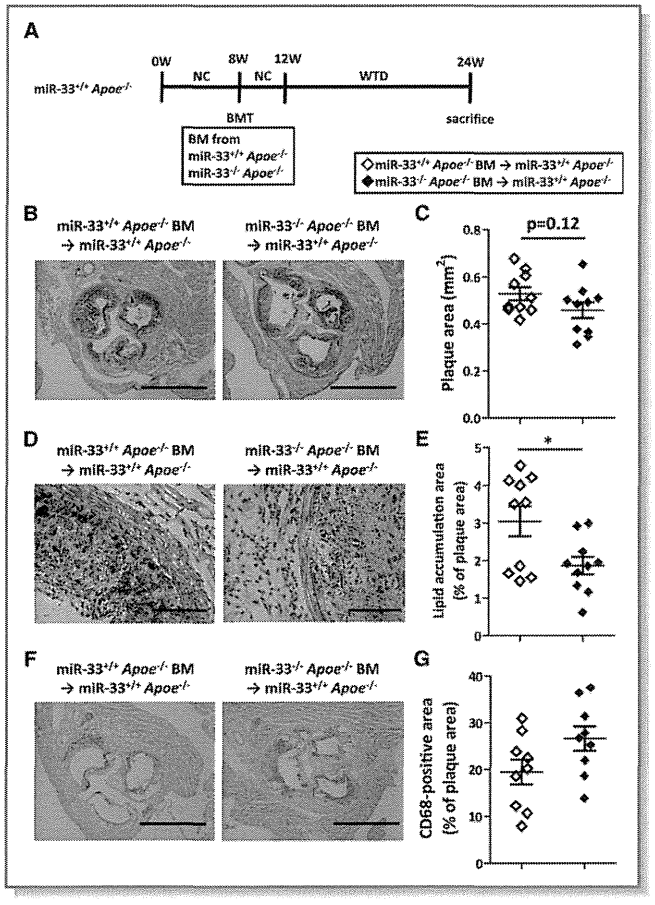

Figure 10. Lipid accumulation area in atherosclerotic lesions was reduced in miR-33$^{+/+}$*Apoe*$^{-/-}$ mice transplanted with miR-33$^{-/-}$ *Apoe*$^{-/-}$ BM. A, Experimental protocol for bone marrow transplantation from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice to miR-33$^{+/+}$*Apoe*$^{-/-}$ mice. B, Representative microscopic images of cross-sections of proximal aorta in mice transplanted with miR-33$^{+/+}$ *Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ BM. Scale bar: 1 mm. C, Quantification of the atherosclerotic plaque area in cross-sections of proximal aorta. Values are mean±SE (n=10 each). D, Representative microscopic images of the lipid accumulation area in atherosclerotic lesions in mice transplanted with miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$ *Apoe*$^{-/-}$ BM. Scale bar: 100 μm. E, Quantification of the lipid accumulation area in cross-sections of proximal aorta. Values are mean±SE (n=10 each); *P<0.05. F, Representative microscopic images of immunohistochemical staining for the macrophage marker CD68 in mice transplanted with miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$ *Apoe*$^{-/-}$ BM. Scale bar: 1 mm. G, Quantification of CD68-positive area in cross-sections of proximal aorta. Values are mean±SE (n=9 each). BM indicates bone marrow.

measured cholesterol efflux capacity in the serum from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. We found that the loss of miR-33 significantly increased the capacity to promote cholesterol efflux, and this may have contributed to the reduction in atherosclerotic plaque volume.

The results of the present BMT experiment revealed that deletion of macrophage miR-33 significantly reduced the lipid content in atherosclerotic plaque. Macrophages are known to ingest apoB-containing lipoproteins and transport ingested lipoprotein-cholesterol from late endosomes to the endoplasmic reticulum (ER) under normal conditions.[41] In advanced lesional macrophages, the accumulation of massive unesterified, or "free," cholesterol is observed, which induces ER-stress-mediated macrophage apoptosis.[42,43] Therefore, we compared atherogenic lipid-induced apoptosis in miR-33$^{+/+}$ *Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ macrophages. Treatment of PEMs in culture with acLDL plus ACAT inhibitor (free cholesterol loading) demonstrated that PEMs from miR-33$^{-/-}$ *Apoe*$^{-/-}$ mice were significantly resistant to apoptosis compared with those from miR-33$^{+/+}$*Apoe*$^{-/-}$ mice. In accordance with this in vitro experiment, apoptotic cells in the lesion area were also reduced in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. It is generally considered that apoptosis of macrophages in advanced plaques may lead to plaque rupture. Thus, the suppression of miR-33 in macrophages may also be beneficial for the prevention of plaque rupture.

To further elucidate the effect of miR-33 deletion on the monocyte/macrophage phenotype, we performed a flow-cytometric analysis of circulating monocytes and a quantitative PCR analysis of RNA from PEMs of miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. We first determined that the total leukocyte count in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice was less than that in miR-33$^{+/+}$*Apoe*$^{-/-}$mice. This was consistent with a previous report that ABCA1, ABCG1, and HDL inhibit the proliferation of hematopoietic stem cells.[44] Because leukocytosis enhances the progression of atherosclerosis, the reduction in leukocytes observed in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice may have had a beneficial effects on atherosclerosis.[45] However, we also detected a higher frequency of Ly6C^{high} monocytes in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice than in miR-33$^{+/+}$ *Apoe*$^{-/-}$ mice, and this could enhance inflammation in atherosclerotic plaque. Tissue macrophages are phenotypically heterogeneous and are broadly characterized according to their activation (polarization) state by the M1/M2 classification system.[46] Some M1 and M2 markers were significantly elevated in miR-33$^{-/-}$*Apoe*$^{-/-}$ PEMs compared with miR-33$^{+/+}$*Apoe*$^{-/-}$ PEMs. RIP140 has been reported to promote the activity of NF-κB and to upregulate the expression of genes implicated in inflammation such as TNFα and IL-6 in macrophages.[38] RIP140 has been shown to be one of the targets of miR-33.[37] Therefore, the enhanced expression of RIP140 under miR-33 deficiency may have affected the expression of IL-6. It is also possible that elevation of M2 markers may indicate the healing process of atherosclerosis in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice and that the phenotypic changes in macrophages may involve feedback mechanisms. In any case, the effect of miR-33 deletion in macrophages is not as simple as a shift from the M1 to the M2 phenotype, as described in a previous report.[19] miR-33 deficiency also

Table 2. Serum Lipid Profiling of MiR-33$^{+/+}$*Apoe*$^{-/-}$ Mice Transplanted With MiR-33$^{+/+}$*Apoe*$^{-/-}$ and MiR-33$^{-/-}$*Apoe*$^{-/-}$ BM by Standard Method

	TC, mg/dL	HDL-C, mg/dL	LDL-C, mg/dL	TG, mg/dL
miR-33$^{+/+}$*Apoe*$^{-/-}$ BM recipient (n=6)	833.5±70.4	12.0±1.4	195.0±15.2	56.3±8.3
miR-33$^{-/-}$*Apoe*$^{-/-}$ BM recipient (n=8)	984.9±54.7	12.1±1.3	209.9±9.0	43.9±6.6
P	NS	NS	NS	NS

Values are mean±SE. BM indicates bone marrow; TC, total cholesterol; TG, triglyceride; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; NS, not significant.

reduced the expression of VCAM-1, which may have influenced atherosclerotic plaque formation.

We also tried to determine the contribution of the loss of miR-33 in recipient mice to atherosclerosis. However, we did not observe a significant elevation of HDL-C level in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice transplanted with BM of miR-33$^{+/+}$*Apoe*$^{-/-}$ or miR-33$^{-/-}$*Apoe*$^{-/-}$ compared with miR-33$^{+/+}$*Apoe*$^{-/-}$ mice transplanted with BM of miR-33$^{+/+}$*Apoe*$^{-/-}$ or miR-33$^{-/-}$*Apoe*$^{-/-}$. Currently, we do not know why HDL-C levels were almost the same in these mice. Experimental conditions such as radiation to the liver and intestine may have reduced the effect of miR-33 deficiency on the increase in HDL-C levels in recipient mice after BMT. Atherosclerotic plaque size was not significantly reduced in miR-33$^{-/-}$*Apoe*$^{-/-}$ recipients compared with the miR-33$^{+/+}$*Apoe*$^{-/-}$ recipients transferred with the same type of BM; this may simply indicate that the rise in HDL-C levels is important in the prevention of atherosclerosis. Our results showed that loss of miR-33 in blood cells reduced the lipid content of atherosclerotic plaque, which may be because of improved cholesterol efflux from macrophages.

A recent report indicated that the inhibition of miR-33a/b in nonhuman primates raised plasma HDL-C and lowered VLDL triglyceride levels.[47] This result was obtained from the administration of antisense miR-33 for a certain period. These data suggest that it may be possible to inhibit miR-33a and miR-33b pharmacologically to raise HDL-C for the treatment of dyslipidemia and atherosclerosis. However, to establish the safety of this therapeutic strategy for the treatment of humans, the complete inhibition of target miRNA and longer-term assessment in animal disease models are required to avoid unexpected side effects.

Moreover, the real targets of miR-33 in vivo can only be clarified by the genetic deletion of miR-33, and the results obtained by antisense oligonucleotide-based medicine may be different from those obtained in miR-33-deficient mice. For example, the administration of miR-21 antagomir prevented pressure-overload-induced cardiac hypertrophy and fibrosis in mice[48]; however, miR-21-deficient mice did not show any cardiac difference with wild-type mice under pressure overload.[49] There also seems to be a substantial difference in the effect of antisense oligonucleotide against miRNA depending on the modification such as LNA based or cholesterol modified, and many additional miRNAs that share an identical or similar sequence may be inactivated. Therefore, caution is needed when interpreting studies that use antagomir approaches to elucidate the function of individual miRNAs in vivo. Previously, we and others showed that miR-33 targets ABCA1 in macrophages and the liver. ABCG1 is another target

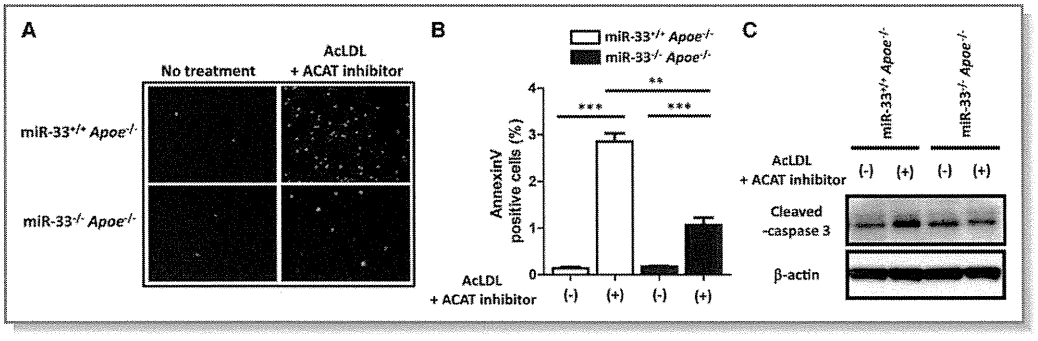

Figure 11. miR-33 deficiency ameliorated free-cholesterol loading-induced macrophage apoptosis. A, Representative microscopic images of Alexa Fluor 448–conjugated annexin V staining. Thioglycollate-elicited peritoneal macrophages from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice were cultured in the presence or absence of acLDL plus ACAT inhibitor for 24 hours. B, Quantification of annexin V–positive macrophages in the presence or absence of acLDL plus ACAT inhibitor for 24 hours. Values are mean±SE; $^{**}P<0.01$, $^{***}P<0.001$. C, Western blotting analysis of cleaved caspase-3 in macrophages in the presence or absence of acLDL plus ACAT inhibitor for 48 hours. β-actin was used as a loading control.

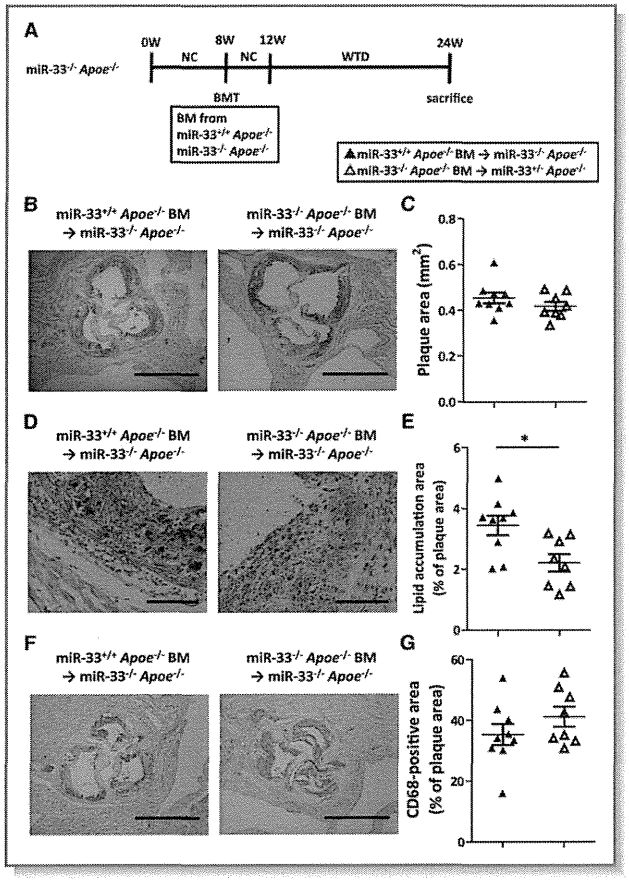

Figure 12. Lipid accumulation area in atherosclerotic lesions was reduced in miR-33$^{-/-}$*Apoe*$^{-/-}$ mice transplanted with miR-33$^{-/-}$*Apoe*$^{-/-}$ BM. A, Experimental protocol for bone marrow transplantation from miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ mice to miR-33$^{-/-}$*Apoe*$^{-/-}$ mice. B, Representative microscopic images of cross-sections of proximal aorta in mice transplanted with miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ BM. Scale bar: 1 mm. C, Quantification of the atherosclerotic lesion area in cross-sections of proximal aorta. Values are mean±SE (n=8 to 9 each). D, Representative microscopic images of the lipid accumulation area in atherosclerotic lesions in mice transplanted with miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ BM. Scale bar: 100 μm. E, Quantification of the lipid accumulation area in cross-sections of proximal aorta. Values are mean±SE (n=8 to 9 each); **P*<0.05. F, Representative microscopic images of immunohistochemical staining for the macrophage marker CD68 in mice transplanted with miR-33$^{+/+}$*Apoe*$^{-/-}$ and miR-33$^{-/-}$*Apoe*$^{-/-}$ BM. Scale bar: 1 mm. G, Quantification of the CD68-positive area in cross-sections of proximal aorta. Values are mean±SE (n=8 to 9 each). BM indicates bone marrow.

in macrophages. In this study, we showed that ABCA1 and CROT in the livers of miR-33$^{-/-}$*Apoe*$^{-/-}$ mice were upregulated compared with those in miR-33$^{+/+}$*Apoe*$^{-/-}$ mice, whereas CPT1a and AMPKα were not. ABCA1, ABCG1, and RIP140 were also upregulated in miR-33$^{-/-}$*Apoe*$^{-/-}$ macrophages. Thus, although many targets have been estimated by many computer algorithms and in vitro experiments such as luciferase-based 3′ UTR analysis and Western blotting, it seems that only some of these actually have any effect in vivo through the chronic complete inhibition of miRNA.

Metabolic syndrome and type 2 diabetes are growing public health concerns worldwide that are associated with complex risk factors for cardiovascular disease (CVD). A previous meta-analysis indicated that a reduction of 2 to 3 mmol/L LDL-C by statins could reduce the risk of CVD by 40% to 50% because there was no evidence of any threshold within the cholesterol range studied.[3] However, this simply means that there is still a substantial risk of CVD despite statins being widely used to lower levels of LDL-C and apolipoprotein B–containing lipoproteins. Major goals in the pursuit of novel therapeutic strategies to target this residual risk have focused on raising the amount and quality of HDL-C to prevent atherosclerosis. Further detailed experiments will be needed to determine whether targeting miR-33 could be a suitable approach for such treatment.

Acknowledgments

We thank Naoya Sowa for providing excellent technical assistance.

Sources of Funding

This work was supported in part by grants for Global COE Program "Center for Frontier Medicine" from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, by grants from the Japan Society for the Promotion of Science, by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (to T. Kimura, T. Kita, K.H., T.H., and K.O.), by a Grant-in-Aid for Scientific Research on Innovative Areas (3307) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to K.O.), by Otsuka

Table 3. Serum Lipid Profiling of MiR-33$^{-/-}$*Apoe*$^{-/-}$ mice transplanted With MiR-33$^{+/+}$*Apoe*$^{-/-}$ and MiR-33$^{-/-}$*Apoe*$^{-/-}$ BM by Standard Method

	TC (mg/dL)	HDL-C (mg/dL)	LDL-C (mg/dL)	TG (mg/dL)
miR-33$^{+/+}$*Apoe*$^{-/-}$ BM recipient (n=9)	875.0±56.0	12.6±1.0	174.6±8.8	108.2±49.0
miR-33$^{-/-}$*Apoe*$^{-/-}$ BM recipient (n=5)	922.8±70.0	13.4±2.1	179.4±12.7	47.2±6.9
P	NS	NS	NS	NS

Values are mean±SE. BM indicates bone marrow; TC, total cholesterol; TG, triglyceride; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; NS, not significant.

Pharmaceutical Co, Ltd – Kyoto University SRP Program (to T. Kimura, T. Kita, M.Y., T.H., and K.O.), by grants from Banyu Life Science Foundation International, Mochida Memorial Foundation, Kimura Memorial Heart Foundation Research Grant for 2010, Japan Heart Foundation Research Grant, Japan Heart Foundation/Novartis Grant for Research Award on Molecular and Cellular Cardiology 2011, and Takeda Memorial Foundation (to T.H.), and by grants from ONO Medical Research Foundation, The Cell Science Research Foundation, Daiichi-Sankyo Foundation of Life Science, and Takeda Memorial Foundation (to K.O.).

Disclosures

None.

References

1. Baigent C, Keech A, Kearney PM, Blackwell L, Buck G, Pollicino C, Kirby A, Sourjina T, Peto R, Collins R, Simes R. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet.* 2005;366:1267–1278.
2. Kearney PM, Blackwell L, Collins R, Keech A, Simes J, Peto R, Armitage J, Baigent C. Efficacy of cholesterol-lowering therapy in 18,686 people with diabetes in 14 randomised trials of statins: a meta-analysis. *Lancet.* 2008; 371:117–125.
3. Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, Bhala N, Peto R, Barnes EH, Keech A, Simes J, Collins R. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet.* 2010;376:1670–1681.
4. Barter P, Gotto AM, LaRosa JC, Maroni J, Szarek M, Grundy SM, Kastelein JJ, Bittner V, Fruchart JC. HDL cholesterol, very low levels of LDL cholesterol, and cardiovascular events. *N Engl J Med.* 2007;357:1301–1310.
5. Wolfram RM, Brewer HB, Xue Z, Satler LF, Pichard AD, Kent KM, Waksman R. Impact of low high-density lipoproteins on in-hospital events and one-year clinical outcomes in patients with non-ST-elevation myocardial infarction acute coronary syndrome treated with drug-eluting stent implantation. *Am J Cardiol.* 2006;98:711–717.
6. Gordon DJ, Probstfield JL, Garrison RJ, Neaton JD, Castelli WP, Knoke JD, Jacobs DR Jr, Bangdiwala S, Tyroler HA. High-density lipoprotein cholesterol and cardiovascular disease. Four prospective American studies. *Circulation.* 1989;79:8–15.
7. Nissen SE, Tsunoda T, Tuzcu EM, Schoenhagen P, Cooper CJ, Yasin M, Eaton GM, Lauer MA, Sheldon WS, Grines CL, Halpern S, Crowe T, Blankenship JC, Kerensky R. Effect of recombinant apoA-I Milano on coronary atherosclerosis in patients with acute coronary syndromes: a randomized controlled trial. *JAMA.* 2003;290:2292–2300.
8. Plump AS, Scott CJ, Breslow JL. Human apolipoprotein A-I gene expression increases high density lipoprotein and suppresses atherosclerosis in the apolipoprotein E-deficient mouse. *Proc Natl Acad Sci USA.* 1994;91:9607–9611.
9. Rubin EM, Krauss RM, Spangler EA, Verstuyft JG, Clift SM. Inhibition of early atherogenesis in transgenic mice by human apolipoprotein AI. *Nature.* 1991;353:265–267.
10. Aiello RJ, Brees D, Bourassa PA, Royer L, Lindsey S, Coskran T, Haghpassand M, Francone OL. Increased atherosclerosis in hyperlipidemic mice with inactivation of ABCA1 in macrophages. *Arterioscler Thromb Vasc Biol.* 2002;22:630–637.
11. Joyce CW, Wagner EM, Basso F, Amar MJ, Freeman LA, Shamburek RD, Knapper CL, Syed J, Wu J, Vaisman BL, Fruchart-Najib J, Billings EM, Paigen B, Remaley AT, Santamarina-Fojo S, Brewer HB Jr. ABCA1 overexpression in the liver of LDLr-KO mice leads to accumulation of pro-atherogenic lipoproteins and enhanced atherosclerosis. *J Biol Chem.* 2006;281:33053–33065.
12. Singaraja RR, Fievet C, Castro G, James ER, Hennuyer N, Clee SM, Bissada N, Choy JC, Fruchart JC, McManus BM, Staels B, Hayden MR. Increased ABCA1 activity protects against atherosclerosis. *J Clin Invest.* 2002;110:35–42.
13. van Eck M, Bos IS, Kaminski WE, Orso E, Rothe G, Twisk J, Bottcher A, Van Amersfoort ES, Christiansen-Weber TA, Fung-Leung WP, Van Berkel TJ, Schmitz G. Leukocyte ABCA1 controls susceptibility to atherosclerosis and macrophage recruitment into tissues. *Proc Natl Acad Sci USA.* 2002;99:6298–6303.
14. Brunham LR, Singaraja RR, Duong M, Timmins JM, Fievet C, Bissada N, Kang MH, Samra A, Fruchart JC, McManus B, Staels B, Parks JS, Hayden MR. Tissue-specific roles of ABCA1 influence susceptibility to atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2009;29:548–554.
15. Tang C, Liu Y, Kessler PS, Vaughan AM, Oram JF. The macrophage cholesterol exporter ABCA1 functions as an anti-inflammatory receptor. *J Biol Chem.* 2009;284:32336–32343.
16. Rayner KJ, Suarez Y, Davalos A, Parathath S, Fitzgerald ML, Tamehiro N, Fisher EA, Moore KJ, Fernandez-Hernando C. MiR-33 contributes to the regulation of cholesterol homeostasis. *Science.* 2010;328:1570–1573.
17. Najafi-Shoushtari SH, Kristo F, Li Y, Shioda T, Cohen DE, Gerszten RE, Naar AM. MicroRNA-33 and the SREBP host genes cooperate to control cholesterol homeostasis. *Science.* 2010;328:1566–1569.
18. Marquart TJ, Allen RM, Ory DS, Baldan A. MiR-33 links SREBP-2 induction to repression of sterol transporters. *Proc Natl Acad Sci USA.* 2010;107:12228–12232.
19. Rayner KJ, Sheedy FJ, Esau CC, Hussain FN, Temel RE, Parathath S, van Gils JM, Rayner AJ, Chang AN, Suarez Y, Fernandez-Hernando C, Fisher EA, Moore KJ. Antagonism of miR-33 in mice promotes reverse cholesterol transport and regression of atherosclerosis. *J Clin Invest.* 2011;121:2921–2931.
20. Plump AS, Smith JD, Hayek T, Aalto-Setala K, Walsh A, Verstuyft JG, Rubin EM, Breslow JL. Severe hypercholesterolemia and atherosclerosis in apolipoprotein E-deficient mice created by homologous recombination in ES cells. *Cell.* 1992;71:343–353.
21. Zhang SH, Reddick RL, Piedrahita JA, Maeda N. Spontaneous hypercholesterolemia and arterial lesions in mice lacking apolipoprotein E. *Science.* 1992;258:468–471.
22. Horie T, Ono K, Horiguchi M, Nishi H, Nakamura T, Nagao K, Kinoshita M, Kuwabara Y, Marusawa H, Iwanaga Y, Hasegawa K, Yokode M, Kimura T, Kita T. MicroRNA-33 encoded by an intron of sterol regulatory element-binding protein 2 (SREBP2) regulates HDL in vivo. *Proc Natl Acad Sci USA.* 2010; 107:17321–17326.
23. Horie T, Ono K, Nishi H, Nagao K, Kinoshita M, Watanabe S, Kuwabara Y, Nakashima Y, Takanabe-Mori R, Nishi E, Hasegawa K, Kita T, Kimura T. Acute doxorubicin cardiotoxicity is associated with miR-146A-induced inhibition of the neuregulin-ERBB pathway. *Cardiovasc Res.* 2010;87:656–664.
24. de la Llera-Moya M, Drazul-Schrader D, Asztalos BF, Cuchel M, Rader DJ, Rothblat GH. The ability to promote efflux via ABCA1 determines the capacity of serum specimens with similar high-density lipoprotein cholesterol to remove cholesterol from macrophages. *Arterioscler Thromb Vasc Biol.* 2010;30:796–801.
25. Khera AV, Cuchel M, de la Llera-Moya M, Rodrigues A, Burke MF, Jafri K, French BC, Phillips JA, Mucksavage ML, Wilensky RL, Mohler ER, Rothblat GH, Rader DJ. Cholesterol efflux capacity, high-density lipoprotein function, and atherosclerosis. *N Engl J Med.* 2011;364:127–135.
26. Asztalos BF, de la Llera-Moya M, Dallal GE, Horvath KV, Schaefer EJ, Rothblat GH. Differential effects of HDL subpopulations on cellular ABCA1- and SR-BI-mediated cholesterol efflux. *J Lipid Res.* 2005;46:2246–2253.
27. Yvan-Charvet L, Ranalletta M, Wang N, Han S, Terasaka N, Li R, Welch C, Tall AR. Combined deficiency of ABCA1 and ABCG1 promotes foam cell accumulation and accelerates atherosclerosis in mice. *J Clin Invest.* 2007;117:3900–3908.
28. Van Eck M, Singaraja RR, Ye D, Hildebrand RB, James ER, Hayden MR, Van Berkel TJ. Macrophage ATP-binding cassette transporter A1 overexpression inhibits atherosclerotic lesion progression in low-density lipoprotein receptor knockout mice. *Arterioscler Thromb Vasc Biol.* 2006;26:929–934.
29. Aparicio-Vergara M, Shiri-Sverdlov R, de Haan G, Hofker MH. Bone marrow transplantation in mice as a tool for studying the role of hematopoietic cells in metabolic and cardiovascular diseases. *Atherosclerosis.* 2010;213:335–344.
30. Paigen B, Morrow A, Holmes PA, Mitchell D, Williams RA. Quantitative assessment of atherosclerotic lesions in mice. *Atherosclerosis.* 1987;68:231–240.
31. Tangirala RK, Rubin EM, Palinski W. Quantitation of atherosclerosis in murine models: correlation between lesions in the aortic origin and in the entire aorta, and differences in the extent of lesions between sexes in LDL receptor-deficient and apolipoprotein E-deficient mice. *J Lipid Res.* 1995;36:2320–2328.
32. Kobayashi T, Tahara Y, Matsumoto M, Iguchi M, Sano H, Murayama T, Arai H, Oida H, Yurugi-Kobayashi T, Yamashita JK, Katagiri H, Majima M, Yokode M, Kita T, Narumiya S. Roles of thromboxane A(2) and prostacyclin in the development of atherosclerosis in apoE-deficient mice. *J Clin Invest.* 2004; 114:784–794.

33. Yao PM, Tabas I. Free cholesterol loading of macrophages is associated with widespread mitochondrial dysfunction and activation of the mitochondrial apoptosis pathway. *J Biol Chem.* 2001;276:42468–42476.

34. Devries-Seimon T, Li Y, Yao PM, Stone E, Wang Y, Davis RJ, Flavell R, Tabas I. Cholesterol-induced macrophage apoptosis requires ER stress pathways and engagement of the type A scavenger receptor. *J Cell Biol.* 2005;171: 61–73.

35. Folch J, Lees M, Sloane Stanley GH. A simple method for the isolation and purification of total lipides from animal tissues. *J Biol Chem.* 1957;226:497–509.

36. Tacke F, Alvarez D, Kaplan TJ, Jakubzick C, Spanbroek R, Llodra J, Garin A, Liu J, Mack M, van Rooijen N, Lira SA, Habenicht AJ, Randolph GJ. Monocyte subsets differentially employ CCR2, CCR5, and CX3CR1 to accumulate within atherosclerotic plaques. *J Clin Invest.* 2007;117:185–194.

37. Ho PC, Chang KC, Chuang YS, Wei LN. Cholesterol regulation of receptor-interacting protein 140 via microRNA-33 in inflammatory cytokine production. *Faseb J.* 2011;25:1758–1766.

38. Fritah A, Christian M, Parker MG. The metabolic coregulator RIP140: an update. *Am J Physiol Endocrinol Metab.* 2010;299:E335–340.

39. Timmins JM, Lee JY, Boudyguina E, Kluckman KD, Brunham LR, Mulya A, Gebre AK, Coutinho JM, Colvin PL, Smith TL, Hayden MR, Maeda N, Parks JS. Targeted inactivation of hepatic ABCA1 causes profound hypoalphalipoproteinemia and kidney hypercatabolism of apoA-I. *J Clin Invest.* 2005;115:1333–1342.

40. Brunham LR, Kruit JK, Pape TD, Parks JS, Kuipers F, Hayden MR. Tissue-specific induction of intestinal ABCA1 expression with a liver × receptor agonist raises plasma HDL cholesterol levels. *Circ Res.* 2006;99:672–674.

41. Scull CM, Tabas I. Mechanisms of ER stress-induced apoptosis in atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2011;31:2792–2797.

42. Thorp E, Li G, Seimon TA, Kuriakose G, Ron D, Tabas I. Reduced apoptosis and plaque necrosis in advanced atherosclerotic lesions of apoe−/− and LDLr−/− mice lacking chop. *Cell Metab.* 2009;9:474–481.

43. Seimon TA, Nadolski MJ, Liao X, Magallon J, Nguyen M, Feric NT, Koschinsky ML, Harkewicz R, Witztum JL, Tsimikas S, Golenbock D, Moore KJ, Tabas I. Atherogenic lipids and lipoproteins trigger CD36-TLR2-dependent apoptosis in macrophages undergoing endoplasmic reticulum stress. *Cell Metab.* 2010;12:467–482.

44. Yvan-Charvet L, Pagler T, Gautier EL, Avagyan S, Siry RL, Han S, Welch CL, Wang N, Randolph GJ, Snoeck HW, Tall AR. ATP-binding cassette transporters and HDL suppress hematopoietic stem cell proliferation. *Science.* 2010;328: 1689–1693.

45. Coller BS. Leukocytosis and ischemic vascular disease morbidity and mortality: is it time to intervene? *Arterioscler Thromb Vasc Biol.* 2005;25: 658–670.

46. Martinez FO, Helming L, Gordon S. Alternative activation of macrophages: an immunologic functional perspective. *Annu Rev Immunol.* 2009;27:451–483.

47. Rayner KJ, Esau CC, Hussain FN, McDaniel AL, Marshall SM, van Gils JM, Ray TD, Sheedy FJ, Goedeke L, Liu X, Khatsenko OG, Kaimal V, Lees CJ, Fernandez-Hernando C, Fisher EA, Temel RE, Moore KJ. Inhibition of miR-33a/b in non-human primates raises plasma HDL and lowers VLDL triglycerides. *Nature.* 2011;478:404–407.

48. Thum T, Gross C, Fiedler J, Fischer T, Kissler S, Bussen M, Galuppo P, Just S, Rottbauer W, Frantz S, Castoldi M, Soutschek J, Koteliansky V, Rosenwald A, Basson MA, Licht JD, Pena JT, Rouhanifard SH, Muckenthaler MU, Tuschl T, Martin GR, Bauersachs J, Engelhardt S. MicroRNA-21 contributes to myocardial disease by stimulating map kinase signalling in fibroblasts. *Nature.* 2008;456:980–984.

49. Patrick DM, Montgomery RL, Qi X, Obad S, Kauppinen S, Hill JA, van Rooij E, Olson EN. Stress-dependent cardiac remodeling occurs in the absence of microRNA-21 in mice. *J Clin Invest.* 2010;120:3912–3916.

BMJ open
accessible medical research

A pragmatic method for electronic medical record-based observational studies: developing an electronic medical records retrieval system for clinical research

Keiichi Yamamoto,[1] Eriko Sumi,[2] Toru Yamazaki,[3] Keita Asai,[3] Masashi Yamori,[3] Satoshi Teramukai,[1] Kazuhisa Bessho,[3] Masayuki Yokode,[2] Masanori Fukushima[4]

To cite: Yamamoto K, Sumi E, Yamazaki T, *et al*. A pragmatic method for electronic medical record-based observational studies: developing an electronic medical records retrieval system for clinical research. *BMJ Open* 2012;**2**:e001622. doi:10.1136/bmjopen-2012-001622

► Prepublication history and additional material for this paper are available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2012-001622).

Received 28 June 2012
Accepted 4 October 2012

For numbered affiliations see end of article.

Correspondence to
Keiichi Yamamoto;
kyamamo@kuhp.kyoto-u.ac.jp

ABSTRACT

Objective: The use of electronic medical record (EMR) data is necessary to improve clinical research efficiency. However, it is not easy to identify patients who meet research eligibility criteria and collect the necessary information from EMRs because the data collection process must integrate various techniques, including the development of a data warehouse and translation of eligibility criteria into computable criteria. This research aimed to demonstrate an electronic medical records retrieval system (ERS) and an example of a hospital-based cohort study that identified both patients and exposure with an ERS. We also evaluated the feasibility and usefulness of the method.

Design: The system was developed and evaluated.

Participants: In total, 800 000 cases of clinical information stored in EMRs at our hospital were used.

Primary and secondary outcome measures: The feasibility and usefulness of the ERS, the method to convert text from eligible criteria to computable criteria, and a confirmation method to increase research data accuracy.

Results: To comprehensively and efficiently collect information from patients participating in clinical research, we developed an ERS. To create the ERS database, we designed a multidimensional data model optimised for patient identification. We also devised practical methods to translate narrative eligibility criteria into computable parameters. We applied the system to an actual hospital-based cohort study performed at our hospital and converted the test results into computable criteria. Based on this information, we identified eligible patients and extracted data necessary for confirmation by our investigators and for statistical analyses with our ERS.

Conclusions: We propose a pragmatic methodology to identify patients from EMRs who meet clinical research eligibility criteria. Our ERS allowed for the efficient collection of information on the eligibility of a given patient, reduced the labour required from the investigators and improved the reliability of the results.

ARTICLE SUMMARY

Article focus

- The focus of this work was to establish a pragmatic methodology to efficiently collect information from electronic medical records (EMRs) about patients who meet clinical research eligibility criteria.

Key messages

- The use of EMR data is necessary to improve clinical research efficiency. However, it is not easy to identify patients who meet research eligibility criteria and collect necessary data from EMRs because the data collection process must integrate various techniques, including the development of a data warehouse and the translation of eligibility criteria into computable criteria. An efficient ERS and a standardised data processing model that integrates these techniques are essential to facilitate clinical research that utilises EMRs.

Strengths and limitations of this study

- Our method uses a specialised data model for patient identification in clinical research and efficient data conversion that does not depend on the EMR database structure when converting narrative criteria to computable criteria.
- We propose that computable criteria should not be a result of the automated conversion of narrative criteria but rather a result of research preparation involving medical concepts that are not expressed logically or explicitly in the narrative criteria. Therefore a large amount of the conversion of the eligibility criteria to computable criteria should be executed at the protocol development stage.
- It is important to further discuss protocol standardisation, including eligibility criteria representation for computable use.
- Enabling medical records retrieval system use in and across multiple institutions is an important future task.

BACKGROUND

Medical information technology has recently advanced in many countries, and enormous amounts of clinical data are already stored as electronic medical records (EMRs). Utilising the data collected in EMRs is necessary to improve clinical research efficiency.[1–3] An EMR is a large database of patient data and is used in observational research to investigate the relationships among diseases, treatments and outcomes,[4–7] to conduct surveillance for rare drug reactions,[4 8] and to recruit patients for clinical trials.[9–13] However, it is not easy to identify patients who meet research eligibility criteria and collect necessary information from EMRs.[2 3] Herein, we describe three major issues concerning EMR-based observational studies: EMR patient data retrieval function, eligibility criteria protocol representation and EMR data accuracy.

To identify patients who meet research eligibility criteria, it is necessary to obtain various types of information stored in EMRs by subject, for example, diagnosis and prescribed medications. However, the EMR database is designed to facilitate online transaction processing for rapid and detail-oriented clinical information searches on individual patients, and the current EMR system does not facilitate this retrieval function.[2 3 14] Data warehouses are essential components of data-driven decision support. To allow for efficient research analyses, EMR data must first be warehoused to enable data analyses across patient populations.[15–21] However, healthcare data modelling is difficult and time consuming because of the complexity of the medical knowledge involved. Thus, the most common approaches to clinical data warehouse modelling are variations on the entity-attribute-value (EAV) model,[22–28] where data are stored in a single table with three columns: entity identification, attribute and attribute value. The EAV design has advantages, including flexibility and ease of storage; however, it requires transforming EAV data into another analytical format before analysis.[25 28] Online analytical processing (OLAP) is most frequently used for searching data stored in the data warehouse.[29–31] OLAP systems in relational databases are typically designed based on Kimball's star schema.[32] However, the star schema was devised to facilitate online measurement analyses. In healthcare, this method can be used to dynamically gather online analyses of numeric data (eg, a specific dose of a drug for a specific disease) in clinical practice. Therefore, this method is not suitable for identifying patients who meet the complicated eligibility criteria for a given clinical research study. Data-modelling methods that facilitate the identification of patients and enable the collection of necessary information from EMRs remain to be established.[28]

Current eligibility criteria are written in a text format that cannot be computationally processed. Additionally, to be applied in actual EMR, eligible criteria need to be integrated with the data model of EMRs.[33] Several investigations have sought to establish computable eligibility criteria.[34–41] However, there is no consensus regarding a standard patient information model,[33] and the eligibility criteria are not yet completely standardised. Using natural language processing (NLP) technologies to convert the text format of eligibility criteria to a computer or to extract patient identifications from EMRs is far from perfect without human intervention.[3 42 43]

Current EMRs have been used to support claims for medical service fees and the treatments administered to each patient; therefore, data gathered specifically for research purposes may be incomplete and unreliable.[2 3 44]

Although various investigations on each technique are executed individually, standardised methods must still be established that integrate these techniques, facilitate the identification of patients who are eligible for clinical research, and collect necessary information from EMRs.

OBJECTIVE

We designed a pragmatic data processing model optimised for patient identification and for the collection of necessary information from EMRs for clinical research. These tools are implemented as an electronic medical records retrieval system (ERS).[44]

This research aimed to demonstrate an ERS and an example of a hospital-based cohort study that used the ERS to identify both patients and exposure. Another aim was to evaluate the feasibility and usefulness of the ERS, the method to convert text form eligible criteria to computable criteria, and a confirmation method to increase research data accuracy.

MATERIALS AND METHODS

Outline of our procedure for patient identification and data collection from the EMR

To identify patients who met the eligibility criteria for the clinical research in question, data were collected in the following ways:

1. The text form of the narrative criteria was converted into computable criteria.
2. A targeted patient list was created.
3. A flag was added for investigators to confirm the targeted patient list.
4. Reports were created for the investigators to confirm.
5. After confirmation by the investigator, the statistical analyses were executed.

EMR retrieval system

In our hospital, EMR use was introduced in 2005; approximately 800 000 cases of clinical information have already been stored. To comprehensively and efficiently collect information about patients participating in clinical research, we developed an ERS.[44]

EMRs store various types of information, integrating billing, pharmacy, radiology, laboratory information and others.[4] In creating the ERS database, we designed a new data model based on the star schema that was

optimised for patient identification in clinical research. We identified nine data categories from EMRs that are useful for clinical research: demographic characteristics, physical findings, diagnostic studies, laboratory tests, diagnoses, progress reports on an EMR template,[44 45] medications and injections, operation records and other treatments. We then designated these categories to 'entities'. In our hospital, the diagnosis is managed by codes that were originally defined by our hospital and mapped with International Statistical Classification of Diseases (ICD) 10 codes[46] for medical insurance purposes. Operations codes were also managed by codes that originally were defined by our hospital and mapped with ICD-9 Clinical Modification codes. We identified available columns (eg, ICD code, diagnosis date) from the EMR data model and designated these columns as 'attributes' of the entities.

Figure 1 presents our data model. In our model, all entities in a given schema are independent and complete; this allows for logical operations and for the creation of eligible patient lists for each respective parameter in a study. The target patient list is generated by combining these patient lists. The data model also supports the inference of medical concepts expressed in the eligibility criteria in reference to corresponding patient data accumulated in EMRs.[33 34]

In our hospital, a replicate of the EMR database known as 'Open DB' was established for the secondary use of accumulated EMR data.[7] A data mart for our ERS was created to ensure that the data retrieval process was practical and independent of the EMR system structure; the data mart was created on the relational database management system by extracting, transforming and loading (ETL) information from the Open DB.[7 44] The ETL process is performed automatically once nightly except for the 'Progress notes by EMR template' entity, which is referred directly from the Open DB to ensure real-time visibility for the eClinical trial.[44]

An OLAP tool was installed to efficiently search through data from multiple patients.[44] The OLAP tool

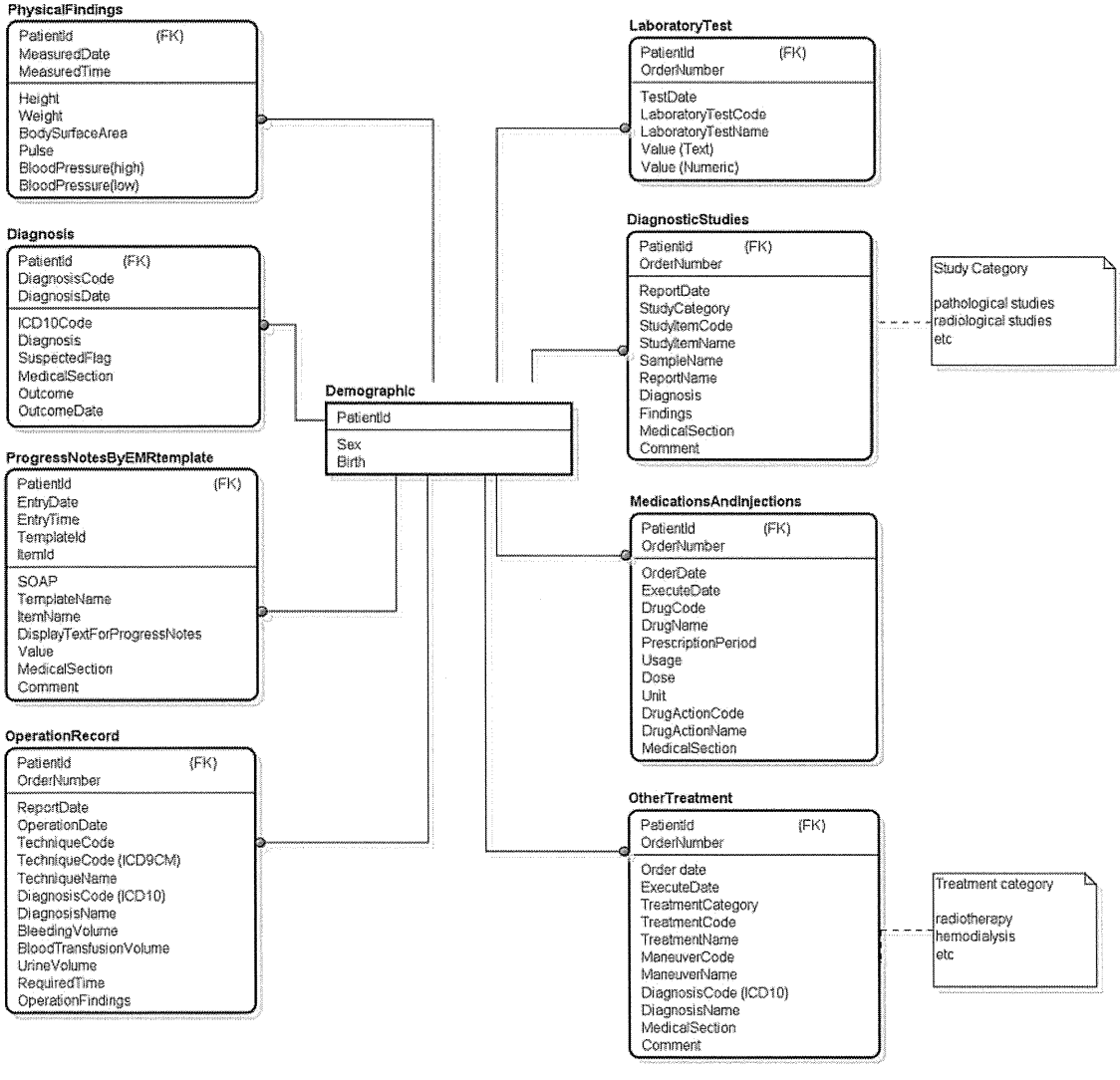

Figure 1 Data model for our electronic medical record retrieval system.

runs in an Internet browser and can generate structured query language (SQL) based on predefined metadata (ie, a data model) by defining logical queries (ie, programmes) using a graphical user interface (GUI). Moreover, this tool allows reports on information retrieved from the browser to be transcribed using hypertext markup language (HTML). The reports are created in various formats, including portable document format (PDF), comma-separated values (CSV) and extensible markup language (XML).[44]

To protect personal information in medical records at our hospital, the EMR network is separated physically from other networks. Our data mart and OLAP servers are deployed in the same EMR network and managed using the same EMR security policies. Additionally, the use of our ERS is limited to clinical research approved by the ethics committee at our hospital, and only designated staff members at our centre are allowed to retrieve data. Our centre creates and manages ERS user identification separate from the EMRs. For the external output of CSV and other data, permission must be obtained from our department of medical informatics, and data extraction must be executed in the presence of supervisors who are responsible for protecting personal information at our hospital.

Application to clinical research

We applied the system to a hospital-based cohort study performed at our hospital titled 'Risk of osteomyelitis of the jaw induced by oral bisphosphonates (BP) in patients taking medications for osteoporosis: a hospital-based cohort study in Japan',[47] in which we identified eligible patients, extracted research data and evaluated the feasibility of our system. The ethics committee at Kyoto University Hospital approved this research. A different paper details the purpose, methods, results and discussion of this research.[47]

This research aimed to estimate the risks for osteomyelitis of the jaw in osteoporosis patients at our hospital who had been exposed to oral BP compared with those who had not.[48 49]

The eligibility criteria were as follows:

Inclusion criteria:

- Patients diagnosed with osteoporosis and treated with osteoporosis medications at Kyoto University Hospital between November 2000 and October 2010.
- Patients aged 20 years or older.

Exclusion criteria:

- Patients with a history of treatment with radiation therapy to the maxillofacial region.
- Patients with primary or metastatic tumours in the maxillofacial region.
- Patients treated with intravenous BP.

The data collected were diagnosis, date of diagnosis, sex, birthdate and the doses and dates when osteoporosis medications, steroids, anticancer drugs, diabetes drugs and HbA1c tests were administered.

Conversion of the text form of the narrative criteria to computable criteria

To identify eligible patients and collect the necessary data from the EMRs, narrative criteria and data must be converted to computable criteria. Such computable criteria include entities, attributes, logical operators (ie, 'and' and 'or'), codes and parameters.[33–37] The clinical research purpose and clinical practice demands made it necessary to perform this task.

We manually executed the conversion from text eligibility criteria to computable criteria. As an example of the conversion from narrative criteria to computable criteria, we present the following two-step conversion procedure:

Step 1: Convert the narrative criteria into entity-level criteria

Medical concepts expressed as narrative criteria are mapped onto entities in the data model and converted into entity-level criteria. This task is manually performed at the protocol development stage of the study by the investigators. For each entity, a criterion is created to extract patients who meet each condition. If exclusive conditions for the same entity must be defined, a different criterion is created. Additionally, the list of codes for drugs and diagnoses (ie, ICD-10) is created, and the period of treatments and others are defined by investigators. In this study, we mapped 'osteoporotic patients' onto two entities (ie, 'diagnosis' and 'medications and injections') and converted it to a combination of two criteria (ie, 'diagnosis of osteoporosis' and 'osteoporosis drug administration'). In the test research, we defined the entity-level criteria according to the entered diagnosis and ordered treatments rather than the diagnostic criteria of the disease. This process reflects that the test research aimed to estimate some risks of osteomyelitis of the jaw with BP administration instead of diagnosing osteoporosis patients accurately. The recorded diagnosis in the EMR was typically designed to ensure payment for medical claims. We thus sought to reduce the number of false-positives by extracting patients with a given treatment type.

Step 2: Convert entity-level criteria into attribute-level criteria (ie, computable criteria)

The abovementioned corresponding codes, date and parameters are mapped onto attributes of the entity-level criteria, and these factors become computable criteria.

Creating a targeted patient list

A targeted patient list is created from the entire set of patients for whom EMRs have been obtained by defining logical queries (ie, programmes defined by the GUI) based on the computable criteria included in the ERS.

Logical queries are first defined in the ERS to identify patients who meet the conditions for each criterion. The ERS automatically generates the SQL necessary for data extraction according to the logical queries. Logical queries are then defined to include or exclude eligible

```
Create View_PatientsList as
Select PatientId From Demographic a
Where
a. PatientId in (
  Select PatientId From Diagnosis
  Where ICD10Code in (osteoporosis ICD10 code list) and
    DiagnosisDate >= '10/01/2000' and DiagnosisDate <= '09/30/2010' and
    SuspectedFlag = 'Fixed' )
and
a. PatientId in (
  Select PatientId From MedicationsAndInjections
  Where DrugCode in (osteoporosis drugs code list) and
    ExecuteDate >= '10/01/2000' and ExecuteDate <= '09/30/2010' )
and
a. PatientId not in (
  Select PatientId From MedicationsAndInjections
  Where DrugCode in (intravenous BP drug code list) and
    ExecuteDate >= '10/01/2000' and ExecuteDate <= '09/30/2010' )
```

Figure 2 Example structured query language (SQL) to create the target patient list.

```
Select PatientId, 'Oral BP administrations', 1 From View_PatientsList a
Where a. PatientId in (
  Select PatientId From MedicationsAndInjections
  Where DrugCode in (oral BP drugs code list) and
    ExecuteDate >= '10/01/2000' and ExecuteDate <= '09/30/2010 ' )
Union all
Select PatientId, 'Oral BP administrations', 0 From View_PatientsList a
Where a. PatientId not in (
  Select PatientId From MedicationsAndInjections
  Where DrugCode in (oral BP drugs code list) and
    ExecuteDate >= '10/01/2000' and ExecuteDate <= '09/30/2010')
Union all
Select PatientId, 'inflammatory jaw condition diagnosis', 1 From View_PatientsList a
Where a. PatientId in (
  Select PatientId From Diagnosis
  Where ICD10Code in (inflammatory conditions of jaws ICD10 code list) and
    DiagnosisDate >= '10/01/2000' and DiagnosisDate <= '09/30/2010' and SuspectedFlag = 'Fixed')
Union all
Select PatientId, 'inflammatory jaw condition diagnosis', 0 From View_PatientsList a
Where a. PatientId not in (
  Select PatientId From Diagnosis
  Where ICD10Code in (inflammatory conditions of jaws ICD10 code list) and
    DiagnosisDate >= '10/01/2000' and DiagnosisDate <= '09/30/2010' and SuspectedFlag = 'Fixed')
```

Figure 3 Example structured query language (SQL) to flag the target patient report for investigator confirmation.

patients who meet each criterion for the demographic entity. The targeted patient list is created by executing the logical query. Figure 2 presents an example of an SQL automatically generated by the ERS.

We thus designed our data model to enable the creation of a targeted patient list by defining the patients extracted from each criterion (ie, 'in' or 'not in') as conditions for the demographic entity that was the unique patient list for the entire hospital. If logical queries are defined using our method, even if the eligibility criteria are complicated, it is not necessary to dramatically change the SQL structure generated in the ERS.

Flagging entries for investigators to confirm

To improve research data accuracy, confirmation by the investigators is necessary. When confirmation is required, additional information is linked.

For the targeted patient list, logical queries are defined to flag certain items according to the investigators' interest. Necessary logical queries are first defined for each criterion. Logical queries are then defined for addition to the patient list as '1' if the data correspond or '0' if they do not. Data sets created by these operations are joined by 'union' and pivoted on a cross-tabulation list using statistical analysis software. We show an example of an SQL generated by the ERS in figure 3.

Create reports for investigators to confirm

To help investigators confirm the targeted patient list, reports are created by linking the findings for diagnostic imaging, pathological diagnosis, operations and other findings. Investigators confirm these entries using the reports and EMR information, including progress notes and images. When the diagnosis history, medication, laboratory results, progress notes and other information are necessary, the same operation is executed for each instance. For example, the list of radiological findings involves 'patient ID', 'study category', 'report name', 'diagnosis', 'findings' and 'comment'. The reports may improve the investigators' confirmation efficiency because they prevent the need to refer to the medical records for each patient who needs confirmation.

Confirmation by the investigator and execution of the statistical analyses

The investigators confirm the accumulated data and execute the statistical analysis. In this test research, two oral and maxillofacial surgeons diagnosed cases by a chart review with an observation of imaging findings.[47]

Systemic evaluation

To evaluate our system, we collected information about the research period using the recall method. For the accuracy of the data collected by the ERS, we evaluated the results after they were confirmed by the investigator.

RESULTS

Computable criteria, datasets and system evaluation

We present the computable criteria in table 1. To increase data accuracy, we collected all of the exclusion criteria for the investigators to confirm. As table 1 shows, we extracted information from EMRs. For investigator confirmation, we also reported all targeted patients using the following lists: osteoporosis drugs administered, oral BP administered, intravenous BP administered, diabetes drugs administered, anticancer drugs administered, steroid drugs administered, osteoporosis diagnoses, oral cancer diagnoses, patients diagnosed with inflammation of the jaw, patients diagnosed with other suspicious diseases, patients diagnosed with diabetes, HbA1c values, radiological findings, pathological findings and radioisotope findings. These data were extracted from the ERS for statistical analyses, presented in CSV format, and analysed using statistics software.

Among the approximately 800 000 cases at our hospital, 8772 were categorised using the terms 'Inclusion

Table 1 Computable criteria for our test research

Criterion	Entity	Operator symbol	Attribute	Operator symbol	Parameter
Created a targeted patient list					
Inclusion criteria: osteoporosis diagnosis	Diagnosis	–	ICD10Code	In	(osteoporosis ICD10 code list)
		and	DiagnosisDate	>=	'10/01/2000'
		and	DiagnosisDate	<=	'09/30/2010'
		and	SuspectedFlag	=	Fixed
Inclusion criteria: osteoporosis drug administrations	Medications and injections	–	DrugCode	in	(osteoporosis drugs code list)
		and	ExecuteDate	>=	'10/01/2000'
		and	ExecuteDate	<=	'09/30/2010'
Added a flag for investigators to confirm the targeted patient list					
Exclusion criteria: oral cancer diagnosis	Diagnosis	–	ICD10Code	in	(oral cancer ICD10 code list)
		and	DiagnosisDate	>=	10/01/2000'
		and	DiagnosisDate	<=	09/30/2010'
		and	SuspectedFlag	=	Fixed
Exclusion criteria: intravenous BP administrations	Medications and injections	–	DrugCode	in	(intravenous BP drugs code list)
		and	ExecuteDate	>=	'10/01/2000'
		and	ExecuteDate	<=	'09/30/2010'
Oral BP administrations	Medications and injections	–	DrugCode	in	(oral BP drugs code list)
		and	ExecuteDate	>=	'10/01/2000'
		and	ExecuteDate	<=	'09/30/2010'
Inflammatory jaw condition diagnosis	Diagnosis	–	ICD10Code	in	(inflammatory conditions of jaws ICD10 code list)
		and	DiagnosisDate	>=	'10/01/2000'
		and	DiagnosisDate	<=	'09/30/2010'
		and	SuspectedFlag	=	Fixed
Other suspicious disease diagnosis	Diagnosis	–	ICD10Code	in	(other suspicious disease ICD10 code list)
		and	DiagnosisDate	>=	'10/01/2000'
		and	DiagnosisDate	<=	'09/30/2010'
		and	SuspectedFlag	=	Fixed
Diabetes diagnosis	Diagnosis	–	ICD10Code	in	(diabetes ICD10 code list)
		and	DiagnosisDate	>=	'10/01/2000'
		and	DiagnosisDate	<=	'09/30/2010'
		and	SuspectedFlag	=	Fixed
Steroid drug administrations	Medications and injections	–	DrugCode	in	(steroid drugs code list)
		and	ExecuteDate	>=	'10/01/2000'
		and	ExecuteDate	<=	'09/30/2010'
Anticancer drug administrations	Medications and injections	–	DrugCode	in	(anticancer drugs code list)
		and	ExecuteDate	>=	'10/01/2000'
		and	ExecuteDate	<=	'09/30/2010'
Diabetes drug administrations	Medications and injections	–	DrugCode	in	(diabetes drugs code list)
		And	ExecuteDate	>=	'10/01/2000'
		And	ExecuteDate	<=	'09/30/2010'
HbA1c test execution	Laboratory test	–	Laboratory TestCode	in	(HbA1c test code)
		and	TestDate	>=	'10/01/2000'
		and	TestDate	<=	'09/30/2010'
Created reports for confirmation by the investigators					
Radiological finding reports	Diagnostic studies	–	ReportName	in	(report name list of oral region)
Pathological finding reports	Diagnostic studies	–	SampleName	contains	'bone'
		Or	SampleName	contains	'jaw'
Radio isotope finding reports	Diagnostic studies	–	–	–	–

BP, bisphosphonates; ID, identifications; ICD, International Classification of Diseases.

criteria: Osteoporosis diagnosis'; among this group, 7195 were further categorised using 'Inclusion criteria: Osteoporosis drug administration'. We then calculated the time that had elapsed since the osteoporosis diagnosis, determined that 7062 patients were aged 20 years or older, and created a targeted patient list. Among those on the targeted patient list, 23 patients were placed under the heading 'Exclusion criteria: Oral cancer diagnosis', 110 under 'Exclusion criteria: Intravenous BP administration', 4200 under 'Oral BP administration', 84 under 'Inflammatory jaw condition diagnosis', 2064 as 'Other suspicious disease diagnosis', 1700 as 'Diabetes diagnosis', 4551 as 'Steroid drug administration', 904 as 'Anticancer drug administrations', 1055 as 'Diabetes drug administrations' and 3641 as 'HbA1c test execution'. Because of the end point considered, patients who were classified under 'Inflammatory jaw condition diagnosis' or 'Other suspicious disease diagnosis' were confirmed using predefined hierarchical diagnostic criteria by investigators who performed the statistical analyses and arranged the research results. We show the schema of data collection and confirmation as figure 4.[47]

The accuracy of the data extracted by the ERS was then characterised. Reviewing the medical records revealed that 2817 patients were not labelled as 'Oral BP administration', including seven (one who received intravenous BP) treated at other hospitals. Six patients had been treated with radiation therapy to the oral and maxillofacial regions. Among the 72 patients classified under 'Inflammatory jaw condition diagnosis', 35 cases and 37 non-cases were identified.

The data extraction period lasted approximately 3 months. Ten meetings were held during the protocol development stage to create and validate the computable criteria and the list of codes for various drugs and diagnoses (ie, ICD-10). The time required for logical query definition when using the ERS was approximately 20 h. The investigator confirmations and statistical analyses took approximately 4 months.

DISCUSSION

We identified eligible patients for this research and extracted the data necessary for confirmation by investigators and for statistical analyses.

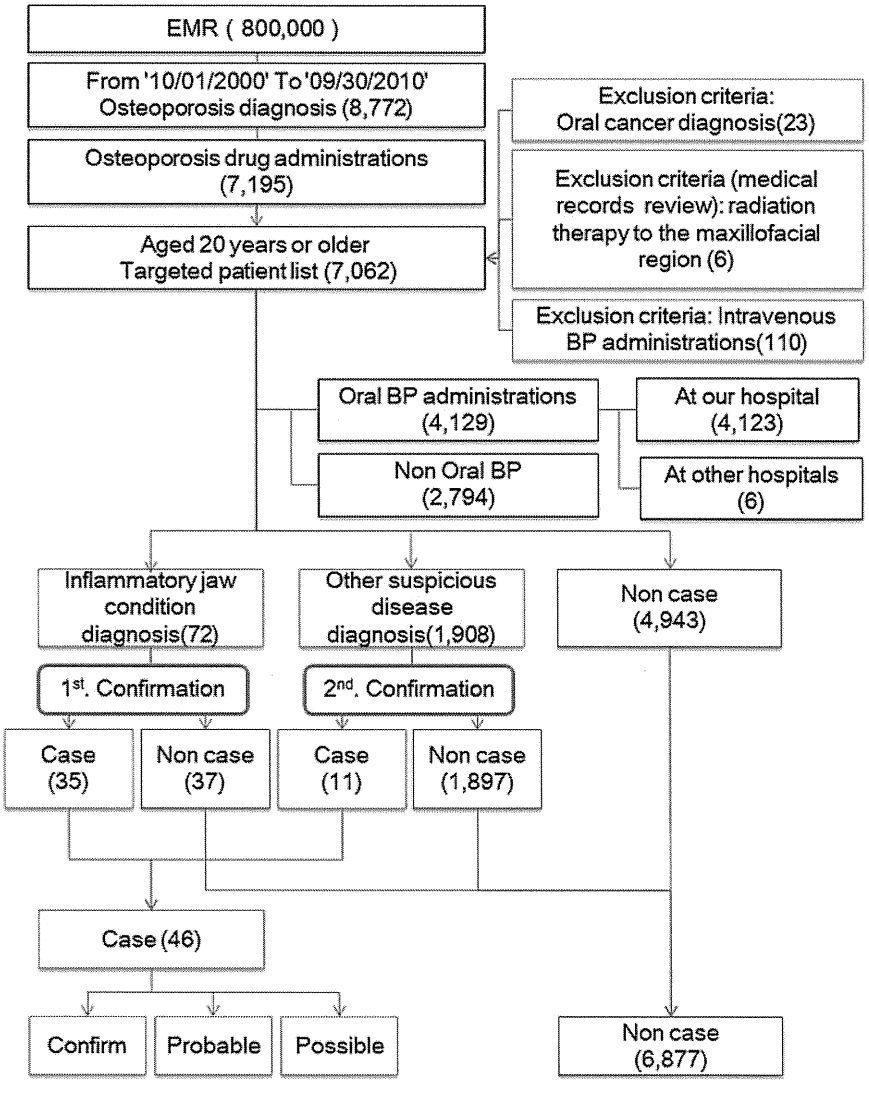

Figure 4 Schema of data collection and confirmation.

We asked the chart reviewers to evaluate the system in a questionnaire about 'the effect of computer programming support for data retrieval from the EMR', 'the result of the data retrieval', 'the positive and negative aspects of our ERS use' and 'the aspects of our method that should be improved'. The investigators evaluating the system mentioned that the following points: (1) the method enabled them to extract the necessary data for diagnosis and drug administration without exception; (2) by screening the entire patient population at the hospital using the ERS, they could identify not just eligible patients in the department of oral and maxillofacial surgery but all eligible patients, which reduced the study bias and (3) by creating reports for confirmation, it enabled investigators to devote their time to reading images, thus effectively reducing the time required for reviewing medical records. The aspects of our method that should be improved are the 'lack of claim data' and the 'administrative complexity of EMR data use'. No negative aspects of our ERS use were noted.

The ERS allowed for the collection of information on patient eligibility by efficiently combining clinical information. Although we did not compare our method with other methods, our proposed method reduced the labour normally required from investigators and improved the reliability of test research results, which indicated that it was useful.

To design the ERS database, we designed a new data model optimised for patient identification. The main differences between our data model and the star schema were as follows: (1) demographic data, which were presented in list form in our EMR system, were presented as a fact-less fact table and (2) date, time, measurements and text information were presented in dimension tables.[32] The most significant characteristic of our method for patient identification is the use of a specialised data model in clinical research and the ability to execute a large number of conversion tasks at the protocol development stage. Data can be converted efficiently in a way that does not depend on the EMR database structure when converting narrative criteria to computable criteria. In this research, we considered whether data were extracted directly from EMRs at the protocol development stage. However, EMR data were recorded in a sequential format for every medical practice, and the database structure was complicated. Comprehending the location and meaning of the necessary data thus required tremendous effort. It was difficult to make precise logical queries for patient identification. However, because our ERS data model was arranged by subjects (eg, tests, diagnosis), it was easy to interpret the available information. Due to the standardisation of computable criteria and SQL possible with the ERS, it was also possible to create computable criteria in little time. Additionally, verifying the patient identification accuracy was easy because it was possible to test each individual criterion.

The SQL generated by our ERS does not reduce the time required for data retrieval. Our ERS also cannot retrieve information that is not in the data model. Current EMRs do not store all necessary data for clinical research, including information related to pregnancy, performance status, cancer stage, availability of transportation to the hospital, specific tests that are not typically performed, drug regimen, outcomes (including death) and adverse events. Additionally, all tests are not administered to all patients, and necessary information may have been recorded in medical records at another hospital.[44] To facilitate EMR use in clinical research, it is necessary to accumulate as much of this information as possible. In the hospital, much of this information does not integrate well with EMRs, including test reports stored only in the departmental system.[50] However, it is important to utilise this information. Additionally, enabling ERS use in and across multiple institutions is also an important future task.

Currently, most clinical research studies that use data from EMRs are planned according to the concept that the primary use of EMRs is for clinical practice and a secondary use is for clinical research.[44] Therefore, most investigators attempt to convert the text form eligibility criteria that already have been defined on a protocol to computable criteria at the data collecting stage.[35 36] However, we propose that computable criteria should not be a result of the automated conversion of narrative criteria but rather a result of research preparation involving medical concepts that are not expressed logically or explicitly in the narrative criteria. Some medical concepts may be interpreted differently depending on the research and the investigator caring for the patients. Additionally, current eligibility criteria are vague or complex, and they do not consider the use of the actual EMR. To convert computable criteria appropriately, high-level medical decisions to answer the research question are required. Therefore, we thought that a large amount of the conversion of the eligibility criteria to computable criteria should be executed at the protocol development stage. In addition, the conversion process should be divided into entity-level conversions that require higher medical decisions and attribute-level conversions. To reduce the burden of conversion, it may be useful to apply NLP technology for the conversion from entity-level criteria to attribute-level criteria. Moreover, it is important to further discuss protocol standardisation, including eligibility criteria representation for computable use. For instance, the attribute-level criteria that describe the search conditions in detail may be useful in global studies to address diseases that vary according to the diagnostic criteria used in each country.

Concerning EMR data accuracy, the ICD10 code (osteomyelitis of the jaw) sensitivity was 48.6% (35/72). The investigators reported six simple diagnosis errors, seven oral BP administrations at other hospitals, and six patients who were treated with radiation therapy in the oral and maxillofacial region.[47] For the accuracy of current EMRs, the investigators had to confirm the information. However, the EMRs provided rich

confirmation data and were useful in improving research data accuracy. In this study, we checked the data from actual EMRs manually and identified patients precisely and extensively using coded information, narrative information, and images. However, only information from existing EMRs was available. Current EMRs have a high degree of flexibility in data entry and are not currently managed for research purposes, which decreases their reliability. It is necessary to improve data quality through quality control without placing too much of a burden on clinical practice. Alternatively, it may be possible to organise data sufficiently before research use.[51–53] Standardising the terminology and exchange formats used in the healthcare setting has facilitated international discourse.[46 54–58] It is necessary to further discuss not only clinical practice but also research purposes, particularly how to utilise various standards when using EMRs beyond the hospital setting.

CONCLUSION

We propose a pragmatic method for EMR-based observational studies. Our ERS is already used to support hospital-based cohort studies, clinical trial recruitment and the eClinical trial infrastructure[44] at our centre. We believe an efficient ERS and standardised data processing model are essential to facilitate clinical research that utilises EMRs.

Author affiliations
[1]Department of Clinical Trial Design and Management, Translational Research Centre, Kyoto University Hospital, Kyoto, Japan
[2]Department of Clinical Innovative Medicine, Translational Research Centre, Kyoto University Hospital, Kyoto, Japan
[3]Department of Oral and Maxillofacial Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan
[4]Translational Research Informatics Centre, Foundation for Biomedical Research and Innovation, Kobe, Japan

Acknowledgements The authors would like to acknowledge the staff of the department of medical informatics of Kyoto University Hospital for their generous support.

Contributors KY designed the study, developed the ERS system, identified the computable eligibility criteria, wrote logical queries, collected data and wrote the manuscript. ES is grant holder who designed the study, developed the ERS system and wrote and edited the manuscript. TY designed and conducted the 'Risk of osteomyelitis of the jaw induced by oral bisphosphonates in patients taking medications for osteoporosis: a hospital-based cohort study in Japan' (BRONJ study) study and the current study, identified the computable eligibility criteria and wrote and edited the manuscript. KA and MY designed and conducted the BRONJ study. ST designed the study and provided comments and feedback. KB is the principal investigator of the BRONJ study. MY owns the ERS system and supervised the study. MF supervised the study and provided comments and feedback. All of the authors read and approved the final manuscript.

Funding This work was supported by the Coordination, Support and Training Program for Translational Research of the Ministry of Education, Culture, Sports, Science and Technology of Japan and by Grants-in-Aid for Scientific Research of Japan (23790566).

Competing interests None.

Ethics approval This research was approved by the ethics committee of Kyoto University Hospital.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No other data are available.

REFERENCES

1. Embi PJ, Payne PR. Clinical research informatics: challenges, opportunities and definition for an emerging domain. *J Am Med Inform Assoc* 2009;16:316–27.
2. Prokosch HU, Ganslandt T. Perspectives for medical informatics. Reusing the electronic medical record for clinical research. *Methods Inf Med* 2009;48:38–44.
3. Wasserman RC. Electronic medical records (EMRs), epidemiology, and epistemology: reflections on EMRs and future pediatric clinical research. *Acad Pediatr* 2011;11:280–7.
4. Dean BB, Lam J, Natoli JL, *et al.* Review: use of electronic medical records for health outcomes research: a literature review. *Med Care Res Rev* 2009;66:611–38.
5. Tannen RL, Weiner MG, Marcus SM. Simulation of the Syst-Eur randomized control trial using a primary care electronic medical record was feasible. *J Clin Epidemiol* 2006;59:25464.
6. Williams JG, Cheung WY, Cohen DR. Can randomised trials rely on existing electronic data? A feasibility study to explore the value of routine data in health technology assessment. *Health Technol Assess* 2003;7: iii, v–x, 1–117.
7. Yamamoto K, Matsumoto S, Tada H, *et al.* A data capture system for outcomes studies that integrates with electronic health records: development and potential uses. *J Med Syst* 2008;32:423–7.
8. Yamamoto K, Matsumoto S, Yanagihara K, *et al.* A data-capture system for post-marketing surveillance of drugs that integrates with hospital electronic health records. *Open Access J Clin Trials* 2011;3:21–6.
9. Embi PJ, Jain A, Clark J, *et al.* Effect of a clinical trial alert system on physician participation in trial recruitment. *Arch Intern Med* 2005;165:2272–7.
10. Campbell MK, Snowdon C, Francis D, *et al.* Recruitment to randomised trials: strategies for trial enrollment and participation study. The STEPS study. *Health Technol Assess* 2007;11:iii, ix-105.
11. Dugas M, Lange M, Müller-Tidow C, *et al.* Routine data from hospital information systems can support patient recruitment for clinical studies. *Clin Trials* 2010;7:183–9.
12. Thadani SR, Weng C, Bigger JT, *et al.* Electronic screening improves efficiency in clinical trial recruitment. *J Am Med Inform Assoc* 2009;16:869–73.
13. Torgerson JS, Arlinger K, Käppi M, *et al.* Principles for enhanced recruitment of subjects in a large clinical trial: the XENDOS study experience. *Control Clin Trials* 2001;22:515–25.
14. Kristianson KJ, Ljunggren H, Gustafsson LL. Data extraction from a semi-structured electronic medical record system for outpatients: a model to facilitate the access and use of data for quality control and research. *Health Inform J* 2009;15:305–19.
15. Shim JP. Past, present, and future of decision support technology. *Decis Support Syst* 2002;33:111–26.
16. Prat N. A UML-based data warehouse design method. *Decis Support Syst* 2006;42:1449–73.
17. Park YT. An empirical investigation of the effects of data warehousing on decision performance. *Inform Manag* 2006;43:51.
18. Schlaps D, Schmid T. Data warehousing in clinical research and development—from clinical data to knowledge portals. *Pharmind* 2004;66:637–46.
19. Grant A, Moshyk A, Diab H, *et al.* Integrating feedback from a clinical data warehouse into practice organisation. *Int J Med Inform* 2006;75:232–9.
20. Junttila K, Meretoja R, Seppälä A, *et al.* Data warehouse approach to nursing management. *J Nurs Manag* 2007;15:155–61.
21. Rubin DL, Desser TS. A data warehouse for integrating radiologic and pathologic data. *J Am Coll Radiol* 2008;5:210–17.
22. Johnson SB. Generic data modeling for clinical repositories. *J Am Med Inform Assoc* 1996;3:328–39.
23. Nadkarni PM, Brandt C. Data extraction and ad hoc query of an entity—attribute—value database. *J Am Med Inform Assoc* 1998;5:511–27.
24. Anhøj J. Generic design of web-based clinical databases. *J Med Internet Res* 2003;5:e27.
25. Chen RS, Nadkarni P, Marenco L, *et al.* Exploring performance issues for a clinical database organized using an entity-attribute-value representation. *J Am Med Inform Assoc* 2000;7:475–87.
26. Dinu V, Nadkarni P. Guidelines for the effective use of entity-attribute-value modeling for biomedical databases. *Int J Med Inform* 2007;76:769–79.