**Supplementary Figure 1. Flow diagram**

First-visit patients to the participating 48 clinics and hospitals, assessed for eligibility at week 0

n=56 261

Switch to mirtazapine

n=284

Based on assessment of depression severity at week 3, remitters continued with sertraline.

Non-remitters were randomized individually (Step 2 randomization)

If unremitted

Assessed (n=111)

Not assessed (imputed) (n=1)

Assessed (n=271)

Not assessed (imputed) (n=2)

Assessed (n=282)

Not assessed (imputed) (n=2)

Assessed (n=274)

Not assessed (imputed) (n=3)

Assessed (n=89)

Not assessed (n=0)

Assessed (n=274)

Not assessed (imputed) (n=6)

Assessed (n=225)

Not assessed (imputed) (n=8)

Assessed (n=229)

Not assessed (imputed) (n=9)

Week 3 PHQ-9 or BDI II not available for 122 (n=71 in 50 mg/d arm, n=51 in 100 mg/d arm). Of those with PHQ-9 at week 3, 23 withdrew consent (n=11 in the 50 mg/d arm, n=12 in the 100 mg/d arm)

Excluded from Step 2 population for protocol violation before Step 2 randomization (n=1 in 100 mg/d arm, subsequently randomized to combination arm); this patient is included for Step 1 analyses

Baseline and week 1 data on PHQ-9 or BDI II not available for 115 (n=61 in 50mg/d arm, n=54 in 100 mg/d arm). Of these, 91 missing baseline only, 15 missing week 1 only, 9 missing both.

Continue with sertraline 50 mg/d

n=112

If remitted

Continue with sertraline 50 mg/d

n=238

If unremitted

Combine sertraline with mirtazapine

n=233

Switch to mirtazapine

n=244

Combine sertraline with mirtazapine

n=273

Continue with sertraline 100 mg/d

n=277

Continue with sertraline 100 mg/d

n=89

If remitted

Intention to titrate sertraline up to 50 mg/d

n= 970

Intention to titrate sertraline up to 100 mg/d

n=1 041

Excluded for being under 25 years of age (n=1) and for exceeding the time limit between week 0 and week 1 (n=1)

Patients with untreated unipolar major depressive episode

n=7 895

Eligible, consented, and registered at the data center through EDC and then cluster-randomized by site at week 1 (Step 1 randomization)

n=2 013

|  |
| --- |
| **Supplementary Table 1. Variable distributions** |
|  |
|  | **(n) imputed** |  | **50 mg/day** |  | **100 mg/day** |  |
|  |  | **Continue** | **Switch** | **Combine** |  |  | **Continue** | **Switch** | **Combine** |  |  |  |
|  |  | **Est** | **(SD)** | **Est** | **(SD)** | **Est** | **(SD)** | **c22** |  | **Est** | **(SD)** | **Est** | **(SD)** | **Est** | **(SD)** | **c22** |  | **c25** |
| **I. Socio-demographics** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Age  | (0) |  | 0.03 | (0.93) | 0.00 | (0.94) | 0.14 | (0.92) | 2.92 |  | -0.07 | (1.0) | -0.04 | (1.0) | -0.03 | (1.0) | 0.14 |  | 7.27 |
| Male sex | (0) |  | 45.0% | (49.4) | 43.4% | (50.0) | 39.5% | (48.85) | 1.54 |  | 49.5% | (49.9) | 57.0% | (48.9) | 55.7% | (49.6) | 3.65 |  | 25.78\* |
| Education |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Continuous | (13) |  | 13.9 | (3.1) | 13.8% | (3.1) | 13.7 | (1.5) | 1.64 |  | 14.2 | (3.3) | 14.2 | (1.7) | 14.0 | (1.7) | 2.72 |  | 13.61\* |
| College graduate | (13) |  | 38.7% | (49.4) | 36.1% | (48.4) | 29.2% | (45.8) | 5.16 |  | 43.3% | (49.9) | 44.4% | (50.6) | 42.5% | (49.6) | 0.20 |  | 18.49\* |
| At least some college | (13) |  | 61.8% | (49.4) | 59.4% | (45.0) | 58.4% | (48.8) | 0.60 |  | 62.8% | (48.3) | 64.8% | (47.2) | 60.1% | (49.6) | 1.32 |  | 3.14 |
| At least high school graduate | (13) |  | 92.4% | (26.2) | 92.2% | (26.6) | 95.3% | (21.4) | 2.56 |  | 93.9% | (23.3) | 94.0% | (23.6) | 91.2% | (28.1) | 1.89 |  | 4.53 |
| Marital status |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Married | (3) |  | 51.3% | (49.4) | 51.6% | (50.0) | 56.2% | (50.4) | 1.45 |  | 48.0% | (49.9) | 50.0% | (50.6) | 53.5% | (49.6) | 1.68 |  | 4.12 |
| Never married | (3) |  | 30.3% | (46.3) | 31.2% | (46.9) | 20.6% | (41.2) | 8.93\* |  | 39.0% | (48.3) | 34.9% | (47.2) | 31.9% | (46.3) | 3.07 |  | 24.76\* |
| Separated/divorced | (3) |  | 16.8% | (37.0) | 15.6% | (35.9) | 19.3% | (39.7) | 1.17 |  | 11.2% | (31.6) | 13.7% | (33.7) | 12.1% | (33.0) | 0.84 |  | 9.10 |
| Widowed  | (3) |  | 1.7% | (12.3) | 1.6% | (12.5) | 3.9% | (19.8) | 2.50 |  | 1.8% | (13.3) | 1.4% | (11.8) | 2.6% | (16.5) | 0.95 |  | 3.56 |
| Employment status |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Employed, full time | (3) |  | 41.6% | (49.4) | 40.2% | (48.4) | 45.9% | (50.4) | 1.73 |  | 35.7% | (48.3) | 38.4% | (48.9) | 37.7% | (47.9) | 0.45 |  | 6.53 |
| Working part time | (3) |  | 11.8% | (32.4) | 8.6% | (28.1) | 10.7% | (30.5) | 1.41 |  | 7.6% | (26.6) | 6.0% | (23.6) | 9.5% | (29.7) | 2.44 |  | 7.48 |
| Employed, on sick leave | (3) |  | 20.6% | (40.1) | 24.6% | (43.7) | 22.3% | (41.2) | 1.10 |  | 31.0% | (46.6) | 33.1% | (47.2) | 29.7% | (46.3) | 0.77 |  | 17.39\* |
| Othera | (3) |  | 26.1% | (43.2) | 26.6% | (43.7) | 21.0% | (41.2) | 2.54 |  | 25.6% | (43.3) | 22.5% | (42.1) | 23.1% | (43.0) | 0.82 |  | 3.48 |
| **II. Physical illness** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Any physical illness | (0) |  | 34.5% | (47.8) | 31.2% | (46.9) | 33.9% | (47.3) | 0.69 |  | 33.2% | (46.6) | 32.8% | (47.2) | 34.4% | (47.9) | 0.19 |  | 0.90 |
| **III. Depression history** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Age-at-onset | (28) |  | 0.07 | (0.93) | -0.03 | (0.94) | 0.08 | (0.92) | 1.89 |  | -0.07 | (1.0) | -0.03 | (1.0) | -0.01 | (1.0) | 0.54 |  | 4.81 |
| Number of depressive episodes | (7) |  | 2.2 | (3.1) | 2.8 | (4.7) | 2.1 | (4.6) | 2.59 |  | 2.6 | (3.3) | 2.2 | (3.4) | 2.1 | (3.3) | 3.06 |  | 5.71 |
| Duration of episode at baseline | (5) |  | -0.03 | (0.77) | 0.07 | (1.4) | 0.15 | (1.5) | 2.84 |  | -0.05 | (0.67) | -0.04 | (0.67) | -0.08 | (0.66) | 0.42 |  | 5.89 |
| **IV. Baseline symptomsb** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| PHQ-9 | (0) |  | -0.12 | (1.1) | -0.04 | (1.1) | -0.01 | (1.1) | 1.51 |  | 0.11 | (1.0) | 0.02 | (1.0) | 0.02 | (1.0) | 1.77 |  | 7.01 |
| BDI-II | (0) |  | -0.05 | (1.1) | -0.08 | (1.1) | -0.17 | (0.92) | 1.68 |  | 0.16 | (1.0) | -0.02 | (1.0) | 0.12 | (1.0) | 6.05\* |  | 20.69\* |
| Factor scores |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Dysphoria | (0) |  | -0.06 | (1.1) | -0.09 | (1.1) | -0.13 | (0.92) | 0.59 |  | 0.18 | (1.0) | 0.02 | (1.0) | 0.05 | (1.0) | 4.24 |  | 16.76\* |
| Negative view of self | (0) |  | -0.06 | (0.93) | -0.08 | (1.1) | -0.20 | (0.92) | 2.74 |  | 0.15 | (1.0) | 0.02 | (1.0) | 0.12 | (1.0) | 2.97 |  | 23.59\* |
| Anhedonia  | (0) |  | 0.00 | (0.93) | 0.01 | (0.94) | 0.02 | (0.92) | 0.07 |  | 0.01 | (1.0) | -0.08 |  | 0.04 | (1.0) | 2.11 |  | 2.33 |
| Suicidality | (0) |  | -0.14 | (1.1) | -0.10 | (0.94) | -0.18 | (1.1) | 0.98 |  | 0.17 | (1.0) | 0.07 | (1.0) | 0.12 | (1.0) | 1.27 |  | 29.46\* |
| Irritability | (0) |  | 0.02 | (0.93) | 0.02 | (0.94) | -0.01 | (0.92) | 0.15 |  | 0.04 | (1.0) | -0.09 | (3.4) | 0.02 | (1.0) | 3.10 |  | 3.46 |
| Sleep/appetite | (0) |  | -0.08 | (1.1) | 0.04 | (0.94) | 0.07 | (1.1) | 3.01 |  | 0.01 | (1.0) | -0.09 | (0.67) | 0.06 | (1.0) | 3.21 |  | 6.58 |
| **V. Week 1 symptomsc** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| PHQ-9 | (0) |  | -0.05 | (1.1) | -0.10 | (1.1) | -0.11 | (1.1) | 0.35 |  | 0.12 | (1.0) | 0.06 | (1.0) | 0.05 | (1.0) | 0.95 |  | 11.19\* |

|  |
| --- |
| **Supplementary Table 1 continued. Variable distributions** |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | **(n) imputed** |  | **50 mg/day** |  | **100 mg/day** |  |  |
|  |  | **Continue** | **Switch** | **Combine** |  |  | **Continue** | **Switch** | **Combine** |  |  |  |
|  |  | **Est** | **(SD)** | **Est** | **(SD)** | **Est** | **(SD)** | **c22** |  | **Est** | **(SD)** | **Est** | **(SD)** | **Est** | **(SD)** | **c22** |  | **c25** |
| BDI-II | (0) |  | -0.07 | (1.1) | -0.08 | (1.1) | -0.15 | (0.92) | 1.08 |  | 0.17 | (1.0) | -0.01 | (1.0) | 0.09 | (1.0) | 4.98 |  | 19.30\* |
| Factor scores |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Dysphoria | (0) |  | -0.72 | (1.2) | -0.70 | (1.2) | -0.77 | (1.2) | 0.47 |  | -0.38 | (1.2) | -0.54 | (1.2) | -0.54 | (1.2) | 3.50 |  | 19.69\* |
| Negative view of self | (0) |  | -0.37 | (1.1) | -0.39 | (1.1) | -0.49 | (1.1) | 1.57 |  | -0.13 | (1.0) | -0.31 | (1.0) | -0.16 | (1.2) | 4.69 |  | 21.82\* |
| Anhedonia  | (0) |  | -0.09 | (1.1) | -0.07 | (1.1) | 0.03 | (1.1) | 1.73 |  | -0.08 | (1.0) | -0.18 | (1.0) | -0.11 | (1.0) | 1.51 |  | 5.51 |
| Suicidality | (0) |  | -0.21 | (1.1) | -0.25 | (1.1) | -0.34 | (0.92) | 2.22 |  | -0.09 | (1.0) | -0.06 | (1.0) | -0.10 | (1.0) | 0.20 |  | 15.96\* |
| Irritability | (0) |  | -0.24 | (1.1) | -0.30 | (0.94) | -0.28 | (0.92) | 0.58 |  | -0.15 | (1.0) | -0.22 | (0.84) | -0.12 | (1.0) | 1.71 |  | 7.60 |
| Sleep/appetite | (0) |  | -0.42 | (1.1) | -0.49 | (1.1) | -0.49 | (1.1) | 0.69 |  | -0.45 | (1.0) | -0.54 | (1.0) | -0.49 | (1.0) | 0.99 |  | 2.01 |
| **VI. Week 3 symptomsc** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| PHQ-9 | (0) |  | -0.06 | (1.1) | 0.00 | (1.1) | -0.06 | (0.92) | 0.54 |  | 0.08 | (1.0) | 0.03 | (1.0) | -0.01 | (1.0) | 1.12 |  | 3.81 |
| BDI-II | (0) |  | -0.05 | (1.1) | 0.00 | (1.1) | -0.06 | (0.92) | 0.43 |  | 0.08 | (1.0) | 0.02 | (1.0) | 0.00 | (1.0) | 1.02 |  | 3.34 |
| Factor scores |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Dysphoria | (0) |  | -1.43 | (1.4) | -1.30 | (1.4) | -1.38 | (1.2) | 1.11 |  | -1.20 | (1.3) | -1.26 | (1.2) | -1.33 | (1.3) | 1.70 |  | 5.34 |
| Negative view of self | (0) |  | -0.70 | (1.1) | -0.70 | (1.2) | -0.79 | (1.1) | 1.04 |  | -0.58 | (1.2) | -0.65 | (1.0) | -0.62 | (1.2) | 0.65 |  | 6.13 |
| Anhedonia  | (0) |  | -0.18 | (1.1) | -0.16 | (1.1) | -0.02 | (1.1) | 3.18 |  | -0.12 | (1.0) | -0.16 | (1.0) | -0.10 | (1.0) | 0.49 |  | 3.69 |
| Suicidality | (0) |  | -0.46 | (0.93) | -0.49 | (0.94) | -0.60 | (0.92) | 3.15 |  | -0.39 | (1.0) | -0.44 | (1.0) | -0.48 | (1.0) | 1.29 |  | 7.72 |
| Irritability | (0) |  | -0.54 | (0.93) | -0.50 | (0.94) | -0.49 | (0.92) | 0.47 |  | -0.50 | (0.83) | -0.52 | (0.8) | -0.52 | (1.0) | 0.10 |  | 0.57 |
| Sleep/appetite | (0) |  | -0.79 | (0.93) | -0.78 | (1.1) | -0.86 | (1.1) | 0.71 |  | -0.84 | (1.0) | -0.87 | (1.0) | -0.88 | (1.0) | 0.21 |  | 2.18 |
| **VII. Symptom changes between baseline and week 3d**   |
| Factor scores |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Dysphoria | (0) |  | -1.4 | (1.2) | -1.21 | (1.2) | -1.25 | (1.2) | 2.06 |  | -1.37 | (1.3) | -1.28 | (1.2) | -1.38 | (1.3) | 1.20 |  | 4.32 |
| Negative view of self | (0) |  | -0.64 | (0.93) | -0.62 | (0.94) | -0.59 | (0.76) | 0.41 |  | -0.73 | (1.0) | -0.67 | (0.84) | -0.74 | (1.0) | 1.26 |  | 6.19 |
| Anhedonia  | (0) |  | -0.18 | (0.93) | -0.18 | (0.94) | -0.04 | (0.76) | 4.25 |  | -0.13 | (0.67) | -0.09 | (0.84) | -0.14 | (0.83) | 0.73 |  | 5.01 |
| Suicidality | (0) |  | -0.33 | (0.93) | -0.38 | (0.94) | -0.42 | (0.92) | 1.38 |  | -0.56 | (0.83) | -0.52 | (1.0) | -0.60 | (1.0) | 1.08 |  | 17.69\* |
| Irritability | (0) |  | -0.57 | (1.1) | -0.52 | (0.94) | -0.48 | (1.1) | 0.85 |  | -0.54 | (1.0) | -0.43 | (0.84) | -0.55 | (1.0) | 2.97 |  | 4.00 |
| Sleep/appetite | (0) |  | -0.72 | (1.1) | -0.82 | (0.94) | -0.93 | (1.1) | 4.38 |  | -0.85 | (1.0) | -0.78 | (0.8) | -0.94 | (1.0) | 2.65 |  | 7.51 |
| Count of number of differences in factor scores (0-6) week 3 versus baseline |  |  |  |  |  |  |  |  |  |  |  |
| Higher (worse sxs) at week 3 | (0) |  | 0.9 | (1.5) | 1.0 | (1.6) | 0.9 | (1.5) | 0.40 |  | 0.9 | (1.7) | 0.9 | (1.7) | 0.8 | (1.7) | 0.36 |  | 1.74 |
| Higher or the same at week 3 | (0) |  | 2.5 | (1.5) | 2.4 | (1.6) | 2.5 | (1.5) | 0.80 |  | 2.3 | (1.7) | 2.4 | (1.7) | 2.2 | (1.7) | 1.36 |  | 6.70 |
| Slightly lower or more at week 3e | (0) |  | 3.9 | (1.5) | 3.8 | (1.6) | 3.9 | (1.5) | 0.18 |  | 3.7 | (1.7) | 3.9 | (1.7) | 3.7 | (1.7) | 5.10 |  | 6.78 |
| **VIII. Side effects at week 3 (dummy variables)f** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Moderate or severe (any duration) | (0) |  | 23.1% | (41.6) | 25.0% | (43.7) | 23.6% | (42.7) | 0.25 |  | 32.9% | (21.6) | 25.4% | (23.6) | 31.9% | (26.4) | 4.66 |  | 11.7\* |
| Moderate or severe 50+% of time | (0) |  | 4.2% | (20.1) | 3.7% | (18.7) | 5.6% | (22.9) | 0.98 |  | 5.1% | (41.6) | 6.0% | (37.1) | 7.3% | (41.3) | 1.22 |  | 4.32 |
| Moderate or severe 90+% of time | (0) |  | 16.8% | (37.0) | 14.8% | (35.9) | 17.6% | (38.2) | 0.77 |  | 22.4% | (33.3) | 16.2% | (30.3) | 21.6% | (34.7) | 4.29 |  | 8.13 |
| Severe (any duration) | (0) |  | 9.7% | (29.3) | 7.0% | (25.0) | 8.6% | (27.5) | 1.19 |  | 12.3% | (30.0) | 9.9% | (27.0) | 13.9% | (29.7) | 2.28 |  | 8.86 |
| Severe 50+% of time | (0) |  | 8.0% | (27.8) | 4.5% | (20.3) | 6.9% | (25.9) | 2.79 |  | 9.7% | (20.0) | 7.7% | (18.5) | 9.5% | (21.5) | 0.87 |  | 8.13 |

|  |
| --- |
| **Supplementary Table 1 continued. Variable distributions** |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | **50 mg/day** |  | **100 mg/day** |  |  |
|  | **(n) imputed** |  | **Continue** | **Switch** | **Combine** |  |  | **Continue** | **Switch** | **Combine** |  |  |  |
|  |  | **Est** | **(SD)** | **Est** | **(SD)** | **Est** | **(SD)** | **c22** |  | **Est** | **(SD)** | **Est** | **(SD)** | **Est** | **(SD)** | **c22** |  | **c25** |
| Severe 90+% of time | (0) |  | 2.5% | (15.4) | 2.0% | (14.1) | 2.6% | (15.3) | 0.18 |  | 4.3% | (41.6) | 3.9% | (40.4) | 4.8% | (43.0) | 0.27 |  | 5.04 |
| Side effects week 3 > week 1  | (0) |  | 16.0% | (37.0) | 14.8% | (35.9) | 16.7% | (38.2) | 0.36 |  | 23.1% | (1.7) | 19.7% | (1.7) | 23.1% | (1.7) | 1.29 |  | 11.10\* |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| (n)  |  |  | (238) | (244) | (233) |  |  | (277) | (284) | (273) |  |  |  |
| (n) |  |  | (715) |  |  | (834) |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Abbreviations. SD, standard deviation; PHQ-9, Patient Health Questionnaire; BDI-II, Beck Depression Inventory 2nd Edition.

\*c2 significant at the 2-sided p<0.05 level.

aHousewife, student, retired, not working.

bAll baseline symptom scales were standardized to a mean of 0 and variance of 1 in the total sample. This was done because these scales have no meaningful metric. Other continuous variables, like years of education and number of prior depressive disorders, were not standardized, in comparison, because they do have meaningful metrics.

cAll week 1 and week 3 symptom scores were standardized to the week 0 means and standard deviations.

dChange scores for continuous variables are in the metric of the standardized week 0 scores.

eThese 3 categories are nested. So the “higher or the same” category includes all those either higher or the same and the “slightly lower or more” includes all those higher as well as those the same and those with negative change scores no less than 0.25 standard deviations (of the change score) below the mean.

fThe dummy variables were selected based on inspection of the week 1, week 3, and weeks 1-3 change distributions of the cross-classification of frequency with severity of side effects.

|  |
| --- |
| **Supplementary Table 2. Algorithms used in the analysis** |
|  |  |  |  |
| **Algorithm** | **Hyperparameters** | **R package, version and authors of the method** | **Description** |
| I. ITR estimation and evaluation  |  | *sg*(Luedtke & van der Laan, 2017) | This doubly-robust framework uses Super Learner to estimate the conditional average treatment effect to build a PTR, and uses a cross-validated targeted maximum likelihood estimator to evaluate the marginal improvement resulting from implementing the PTR in a population. |
|  |
|  |  |  |  |
| II. Super Learner  |  | *SuperLearner v2.0-26* (van der Laan, Polley, & Hubbard, 2007) | Super Learner is an ensemble machine learning approach that uses cross-validation (CV) to select a weighted combination of predicted outcome scores across a collection of candidate algorithms (learners) to yield an optimal combination according to a pre-specified criterion that performs at least as well as the best component algorithm. |
|  |  |  |  |
| III. Learners in the Super Learner library |  |  |  |
|  |  |  |  |
|  a. Meana | NA | *Stats v3.6.1* (Nelder & Wedderburn, 1972) | Arithmetic mean |
| b. Generalized linear models |  |  |
|  GLM – *SL.glm*a | NA | Maximum likelihood estimation with flexible link function. |
| GLM with interaction – *SL.glm.interaction* | Default | GLM with two-way interactions. |
| Stepwise GLM – *SL.step* | Default | GLM with stepwise variable selection. |
|  |  |  |  |
|  |  |  |  |
| c. Elastic Net – *SL.glmnet* | alpha = (0, 0.5, 1) | *Glmnet v3.0-1* (Friedman, Hastie, & Tibshirani, 2010) | Elastic net is a regularization method that minimizes the problem of overlap among predictors by explicitly penalizing over-fitting with a composite penalty λ{MPP x Plasso + (1- MPP) X Pridge}, where MPP is a mixing parameter penalty with values between 0 and 1 that controls relative weighting between the lasso penalty (Plasso) and the ridge penalty (Pridge). The parameter λ controls the total amount of penalization. The ridge penalty handles multicollinearity by shrinking all coefficients smoothly towards 0 but retains all variables in the model. The lasso penalty allows simultaneous coefficient shrinkage and variable selection, tending to select at most one predictor in each strongly correlated set, but at the expense of giving unstable estimates in the presence of high multicollinearity. The elastic net approach of combining the ridge and lasso penalties has the advantage of yielding more stable and accurate estimates than either ridge or lasso alone while maintaining model parsimony.  |
|  |  |  |  |
| d. Spline |  |  |  |
|  |  |  |  |
| Adaptive polynomial splines – *SL.polymars* | gcv = (4, 5, 6) | *polspline v1.1.17* (Kooperberg, 2015) | Adaptive polynomial splines are like adaptive splines but differ in the order in which basis functions (e.g., linear versus nonlinear) are added to build the final model. |
| e. Decision trees – bagging |  |  |  |
| Random Forest – *SL.ranger* | min.node.size = (10, 12, 14, 16) | *ranger v0.11.2* (Wright & Ziegler, 2017) | Independent variables are partitioned (based on contiguous values) and stacked to build short decision trees that are combined (ensemble) to create an aggregate “forest”. Random forest builds  |
| **Supplementary Table 2 continued. Algorithms used in the analysis**  |
|  |  |  |  |
| **Algorithm** | **Hyperparameters** | **R package, version and authors of the method** | **Description** |
|  |  |  | numerous trees in bootstrapped samples and generates an aggregate tree by averaging across trees, thereby reducing over-fitting. |
|  f. Decision trees - boosting  |  |  |  |
| Gradient Boosting Machine – *SL.gbm* | Default | *gbm v2.1.5* (Friedman, 2002) | GBMs build a sequential ensemble of shallow successive regression trees that iteratively learn the residuals from prior trees. This is a flexible method, where the number of trees, interaction depth, and shrinkage are leveraged to build flexible models.  |
| Extreme Gradient Boosting – *SL.xgboost* | max\_depth = (4,5,6) | *xgboost v0.90.0.2* (Chen, & Guestrin, 2016) | A fast and efficient implementation of gradient boosting. |
|  |  |  |  |
| g. Support Vector Machine – *SL.ksvm* | kernel = (linear, polynomial, radial); C = (1, 1.5, 2) | *kernlab v0.9-29* (Scholkopf, Platt, & Shawe-Taylor, 2001) | Classification algorithm that aims to create maximum distance between classes of outcome separated by an optimally selected hyperplane (linear, polynomial or radial kernel). |
|  |  |  |  |
| h. BayesGLM – *SL.bayesglm* | Default | *arm v1.10-1* (Gelman, Jakulin, & Pittau, 2008) | Bayesian functions for generalized linear modeling with independent normal, t, or Cauchy prior distributions for the coefficients. The framework takes a generalized linear model approach and uses an approximate EM (expectation-maximization) algorithm to update the betas iteratively.  |
|  |  |  |  |
| i. Generalized Additive Models – *SL.gam* | Default | *gam v1.16.1* (Hastie & Tibshirani, 1986) | An additive modeling technique where the impact of predictors is captured through smooth functions that can be nonlinear. |
|  |  |  |  |

aAlgorithms only used for covariate balance.

|  |
| --- |
| **Supplementary Table 3. Variable distributions by optimal treatment for second-line 9-week remission**  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | **50 mg/day** |  |  | **100 mg/day** |  |  |
|  | **Continue** | **Switch** | **Combine** |  |  | **Continue** | **Switch** | **Continue** |  |  |  |
|  | **Est** | **(SD)** | **Est** | **(SD)** | **Est** | **(SD)** | **c22** |  | **Est** | **(SD)** | **Est**  | **(SD)** | **Est** | **(SD)** | **c22** |  | **c25** |
| **I. Socio-demographics** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Age | -0.72 | (0.60) | 0.63 | (0.88) | -0.02 | (0.94) | 171.86\* |  | 0.17 | (1.05) | 0.16 | (1.14) | -0.15 | (0.95) | 16.23\* |  | 69.81\* |
| Male sex | 0.58a | (0.50) | 0.48 | (0.50) | 0.39 | (0.49) | 10.02\* |  | 0.79a | (0.41) | 0.43 | (0.50) | 0.58 | (0.49) | 29.00\* |  | 14.63\* |
| Education |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Continuous | 14.03a | (2.59) | 13.24 | (2.23) | 13.94 | (2.13) | 12.21\* |  | 14.65a | (2.59) | 14.03 | (2.39) | 14.16 | (2.41) | 1.87 |  | 6.67\* |
| College graduate | 42.2%a | (49.8) | 27.2% | (44.6) | 36.0% | (48.0) | 6.13\* |  | 52.9%a | (50.7) | 38.5% | (48.8) | 45.0% | (49.8) | 4.22 |  | 7.20\* |
| At least some college | 62.5% | (48.8) | 48.3% | (50.1) | 63.0% | (48.3) | 10.27\* |  | 70.6% | (46.2) | 60.7% | (49.0) | 62.9% | (48.3) | 1.45 |  | 6.53\* |
| At least high school graduate | 89.1% | (31.5) | 88.1% | (32.5) | 95.4% | (21.0) | 8.71\* |  | 94.1% | (23.9) | 94.3% | (23.3) | 92.4% | (26.4) | 1.01 |  | 2.30 |
| Marital status |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Married | 28.1% | (45.3) | 53.6% | (50.0) | 56.0% | (49.7) | 21.29\* |  | 76.5 | (43.1) | 46.7% | (50.0) | 50.5% | (50.0) | 14.05\* |  | 3.55 |
| Never married | 48.4% | (50.4) | 10.6% | (30.9) | 29.8% | (45.8) | 51.73\* |  | 17.6% | (38.7) | 33.6% | (47.3) | 37.1% | (48.3) | 8.22\* |  | 13.17\* |
| Separated/divorced | 23.4% | (42.7) | 33.8%a | (47.5) | 11.4% | (31.8) | 32.54\* |  | 5.9% | (24.9) | 16.4%a | (37.1) | 11.0% | (31.3) | 6.27\* |  | 26.98\* |
| Widowed  | 0.00 | (0.00) | 0.02 | (0.14) | 0.03 | (0.17) | 17.39\* |  | 0.00 | (0.00) | 0.03 | (0.18) | 0.01 | (0.12) | 16.33\* |  | 33.72\* |
| Employment status |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Employed, full time | 43.8% | (50.0) | 36.4% | (48.3) | 44.2% | (49.7) | 3.02 |  | 52.9% | (50.7) | 37.3% | (48.5) | 36.3% | (48.1) | 3.55 |  | 3.38 |
| Working part time | 9.4% | (29.4) | 13.9% | (34.7) | 9.4% | (29.2) | 2.15 |  | 5.9% | (23.9) | 8.6% | (28.1) | 7.4% | (26.2) | 0.53 |  | 1.75 |
| Employed, on sick leave | 21.9% | (41.7) | 21.2% | (41.0) | 23.0% | (42.4) | 0.24 |  | 32.4% | (47.5) | 27.0% | (44.5) | 33.1% | (47.1) | 3.04 |  | 2.00 |
| Otherb | 25.0% | (43.6) | 28.5% | (45.3) | 23.4% | (42.4) | 1.51 |  | 8.8% | (28.8) | 27.0% | (44.5) | 23.2% | (42.2) | 10.49\* |  | 4.17 |
| **II. Physical illness** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Any physical illness | 25.0% | (43.6) | 39.7%a | (49.1) | 32.2% | (46.8) | 5.17 |  | 32.4% | (47.5) | 38.5a | (49) | 31 | (46) | 3.85 |  | 8.17\* |
| **III. Depression history** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Age-at-onset | -0.80 | (0.74) | 0.63 | (0.95) | -0.03 | (0.92) | 143.55\* |  | 0.29 | (1.08) | 0.13 | (1.13) | -0.13 | (0.93) | 13.61\* |  | 58.41\* |
| Number of depressive episodes | 2.75 | (4.55) | 3.23 | (7.27) | 2.06 | (2.65) | 5.01 |  | 1.29 | (0.52) | 2.09 | (2.35) | 2.44 | (3.42) | 53.80\* |  | 1.05 |
| Duration of episode at baseline | 0.42 | (2.34) | -0.04 | (0.59) | 0.05 | (1.24) | 3.52 |  | -0.02 | (0.45) | -0.06 | (0.84) | -0.05 | (0.58) | 0.25 |  | 3.37 |
| **IV. Baseline symptomsc** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| PHQ-9 | 0.06 | (1.02) | -0.18d | (1.05) | -0.03 | (1.06) | 3.14 |  | -0.06 | (0.94) | -0.31d | (0.96) | 0.21 | (0.90) | 50.87\* |  | 35.79\* |
| BDI-II | 0.50 | (1.07) | -0.09 | (0.99) | -0.18 | (1.01) | 23.14\* |  | -0.23 | (0.99) | -0.55 | (0.89) | 0.38 | (0.86) | 192.86\* |  | 79.71\* |
| Factor scores |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Dysphoria | 0.28 | (1.04) | -0.14d | (1.04) | -0.13 | (1.00) | 9.33\* |  | -0.10 | (0.98) | -0.30d | (1.00) | 0.26 | (0.92) | 56.10\* |  | 29.75\* |
|  Negative view of self | 0.45 | (0.93) | -0.17 | (1.00) | -0.17 | (0.98) | 25.85\* |  | -0.27 | (1.15) | -0.61 | (0.96) | 0.43 | (0.81) | 222.50\* |  | 106.11\* |
|  Anhedonia | 0.00 | (0.94) | 0.16 | (1.04) | -0.03 | (1.03) | 3.91 |  | -0.34 | (0.80) | -0.05 | (1.07) | 0.03 | (0.95) | 7.15\* |  | 2.01 |
| **Supplementary Table 3 continued. Variable distributions by optimal treatment for second-line 9-week remission**  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | **50 mg/day** |  | **100 mg/day** |  |  |
|  | **Continue** | **Switch** | **Combine** |  |  | **Continue** | **Switch** | **Continue** |  |  |  |
|  | **Est** | **(SD)** | **Est** | **(SD)** | **Est** | **(SD)** | **c22** |  | **Est** | **(SD)** | **Est**  | **(SD)** | **Est** | **(SD)** | **c22** |  | **c25** |
|  Irritability | 0.46a | (1.05) | 0.06 | (0.97) | -0.06 | (1.01) | 14.67\* |  | 0.28a | (1.00) | -0.51 | (0.86) | 0.19 | (0.96) | 107.36\* |  | 57.04\* |
|  Sleep/appetite | -0.04 | (1.03) | 0.10 | (0.96) | -0.01 | (0.99) | 1.69 |  | 0.09 | (1.13) | -0.25 | (0.98) | 0.09 | (1.00) | 20.20\* |  | 7.37\* |
| **V. Week 1 symptomse** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| PHQ-9 | 0.64a | (0.95) | -0.04 | (1.02) | -0.19 | (1.03) | 42.62\* |  | 0.30a | (1.05) | -0.32 | (0.97) | 0.23 | (0.89) | 58.25\* |  | 45.48\* |
| BDI-II | 0.69 | (1.16) | -0.02 | (1.01) | -0.22 | (0.97) | 38.93\* |  | 0.11 | (1.00) | -0.48 | (0.90) | 0.33 | (0.89) | 138.24\* |  | 62.33\* |
| Factor scores |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Dysphoria | 0.14a | (1.32) | -0.64 | (1.29) | -0.87 | (1.17) | 35.82\* |  | -0.16a | (0.96) | -0.95 | (1.14) | -0.31 | (1.11) | 58.35\* |  | 41.70\* |
|  Negative view of self | 0.37 | (1.10) | -0.44 | (1.05) | -0.51 | (1.04) | 37.86\* |  | -0.41 | (1.29) | -0.90 | (1.05) | 0.12 | (0.91) | 174.41\* |  | 86.91\* |
|  Anhedonia | 0.03 | (0.99) | 0.08 | (1.05) | -0.09 | (1.08) | 3.34 |  | -0.20 | (0.87) | -0.21 | (1.08) | -0.08 | (0.98) | 2.71 |  | 0.18 |
|  Suicidality  | 0.30 | (1.29) | -0.34d | (1.02) | -0.32 | (0.98) | 14.63\* |  | -0.15 | (1.23) | -0.45d | (0.88) | 0.08 | (0.98) | 58.22\* |  | 36.10\* |
|  Irritability | 0.28a | (1.16) | -0.18 | (0.94) | -0.37 | (0.89) | 21.74\* |  | 0.22a | (0.94) | -0.69 | (0.73) | 0.04 | (0.91) | 155.84\* |  | 66.61\* |
|  Sleep/appetite | -0.01a | (1.12) | -0.25 | (1.08) | -0.59 | (1.03) | 23.07\* |  | -0.27a | (1.25) | -0.65 | (1.02) | -0.44 | (1.01) | 8.31\* |  | 11.07\* |
| **VI. Week 3 symptomse** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| PHQ-9 | 0.88a | (0.98) | 0.13 | (1.03) | -0.21 | (0.97) | 75.97\* |  | 0.48a | (0.97) | -0.14 | (0.99) | 0.09 | (0.95) | 16.60\* |  | 58.98\* |
| BDI-II | 0.88a | (1.07) | 0.12 | (1.06) | -0.20 | (0.92) | 66.44\* |  | 0.38a | (0.91) | -0.32 | (1.00) | 0.17 | (0.95) | 47.46\* |  | 54.60\* |
| Factor scores |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Dysphoria | -0.35a | (1.33) | -1.12 | (1.36) | -1.58 | (1.22) | 57.22\* |  | -0.52a | (1.31) | -1.46 | (1.33) | -1.22 | (1.20) | 17.41\* |  | 50.45\* |
|  Negative view of self | 0.23 | (1.16) | -0.71 | (1.13) | -0.86 | (1.06) | 51.97\* |  | -0.44 | (1.14) | -1.17 | (1.13) | -0.38 | (1.05) | 85.81\* |  | 65.41\* |
|  Anhedonia | -0.02 | (1.11) | 0.03 | (1.07) | -0.18 | (1.06) | 5.39 |  | -0.05 | (0.87) | -0.14 | (1.12) | -0.13 | (0.99) | 0.28 |  | 5.67 |
|  Suicidality  | 0.07 | (1.29) | -0.49d | (0.90) | -0.60 | (0.89) | 17.43\* |  | -0.33 | (1.11) | -0.66d | (0.86) | -0.34 | (1.00) | 21.38\* |  | 18.13\* |
|  Irritability | 0.29a | (1.13) | -0.36 | (0.94) | -0.66 | (0.81) | 51.84\* |  | -0.09a | (0.76) | -0.82 | (0.81) | -0.41 | (0.88) | 55.12\* |  | 51.18\* |
|  Sleep/appetite | 0.10a | (1.03) | -0.54 | (1.04) | -1.00 | (0.96) | 79.69\* |  | -0.68a | (1.09) | -0.93 | (1.00) | -0.85 | (0.99) | 2.21 |  | 43.70\* |
| **VII. Symptom changes between baseline and week 3f** |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Factor scores |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Dysphoria | -0.63a | (1.15) | -0.97 | (1.28) | -1.45d | (1.22) | 38.49\* |  | -0.42a | (1.35) | -1.17 | (1.22) | -1.48d | (1.19) | 28.66\* |  | 68.79\* |
|  Negative view of self | -0.22a | (0.87) | -0.54 | (0.88) | -0.69d | (0.84) | 18.69\* |  | -0.17a | (0.62) | -0.56 | (0.84) | -0.81d | (0.89) | 41.84\* |  | 62.64\* |
|  Anhedonia | -0.02a | (0.94) | -0.12 | (0.80) | -0.15d | (0.83) | 1.24 |  | 0.29a | (0.61) | -0.09 | (0.91) | -0.16d | (0.76) | 16.93\* |  | 18.46\* |
|  Suicidality  | -0.23 | (1.02) | -0.19 | (0.83) | -0.45 | (0.88) | 12.17\* |  | -0.09 | (0.88) | -0.42 | (0.87) | -0.65 | (0.96) | 21.36\* |  | 47.16\* |
|  Irritability | -0.16 | (1.24) | -0.42 | (1.01) | -0.60 | (0.97) | 10.09\* |  | -0.37 | (1.00) | -0.32 | (0.92) | -0.60 | (0.93) | 16.45\* |  | 26.96\* |
|  Sleep/appetite | 0.14 | (1.10) | -0.65 | (1.10) | -1.00 | (1.06) | 66.81\* |  | -0.77 | (0.97) | -0.68 | (1.13) | -0.94 | (1.17) | 8.69\* |  | 55.91\* |
| Count of number of differences in factor scores (0-6) week 3 versus baseline |  |  |  |  |  |  |  |  |  |  |  |  |
| **Supplementary Table 3 continued. Variable distributions by optimal treatment for second-line 9-week remission** |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | **50 mg/day** |  | **100 mg/day** |  |  |
|  | **Continue** | **Switch** | **Combine** |  |  | **Continue** | **Switch** | **Continue** |  |  |  |
|  | **Est** | **(SD)** | **Est** | **(SD)** | **Est** | **(SD)** | **c22** |  | **Est** | **(SD)** | **Est**  | **(SD)** | **Est** | **(SD)** | **c22** |  | **c25** |
| Higher (worse sxs) at week 3 | 1.95a | (1.25) | 1.16 | (1.27) | 0.74d | (1.06) | 64.07\* |  | 1.56a | (1.50) | 1.02 | (1.25) | 0.77d | (1.03) | 15.69\* |  | 72.03\* |
| Higher or the same at week 3 | 3.31a | (1.57) | 2.85 | (1.51) | 2.26d | (1.37) | 39.27\* |  | 3.03a | (1.36) | 2.64 | (1.46) | 2.14d | (1.41) | 29.91\* |  | 68.90\* |
| Slightly lower at week 3e | 4.39a | (0.85) | 4.07 | (1.13) | 3.74d | (1.32) | 31.47\* |  | 4.41a | (0.86) | 4.10 | (1.08) | 3.61d | (1.34) | 46.15\* |  | 76.85\* |
| **VIII. Side effects at week 3 (dummy variables)h** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Moderate or severe (any duration) | 0.31 | (0.47) | 0.25 | (0.44) | 0.23 | (0.42) | 2.20 |  | 0.15 | (0.36) | 0.29 | (0.46) | 0.31 | (0.46) | 6.76\* |  | 0.18 |
| Moderate or severe 50+% of time | 0.23 | (0.43) | 0.15 | (0.36) | 0.16 | (0.37) | 2.03 |  | 0.15 | (0.36) | 0.19 | (0.39) | 0.21 | (0.41) | 1.21 |  | 0.48 |
| Moderate or severe 90+% of time | 0.05 | (0.21) | 0.03 | (0.18) | 0.05 | (0.21) | 0.74 |  | 0.03 | (0.17) | 0.09 | (0.28) | 0.05 | (0.22) | 3.78 |  | 1.60 |
| Severe (any duration) | 0.19 | (0.39) | 0.07 | (0.26) | 0.07 | (0.26) | 5.19 |  | 0.03 | (0.17) | 0.11 | (0.31) | 0.13 | (0.34) | 9.58\* |  | 0.96 |
| Severe 50+% of time | 0.16 | (0.37) | 0.04 | (0.20) | 0.06 | (0.24) | 6.01\* |  | 0.03 | (0.17) | 0.09 | (0.29) | 0.09 | (0.29) | 4.22 |  | 1.45 |
| Severe 90+% of time | 0.03 | (0.18) | 0.01 | (0.11) | 0.03 | (0.16) | 1.38 |  | 0.03 | (0.17) | 0.05 | (0.23) | 0.04 | (0.20) | 0.89 |  | 0.23 |
| Side effects week 3 > week 1  | 0.25 | (0.44) | 0.20 | (0.40) | 0.13 | (0.34) | 6.71\* |  | 0.06 | (0.24) | 0.21 | (0.41) | 0.23 | (0.42) | 15.66\* |  | 0.65 |
|  | (64) | (151) | (500) |  |  | (34) | (244) | (556) |  |  |
|  | (715) |  |  | (834) |  |

Abbreviations. SE, standard error; PHQ-9, Patient Health Questionnaire; BDI-II, Beck Depression Inventory 2nd Edition.

\*c2 significant at the 2-sided p<0.05 level.

aConsistently highest value across samples.

bHousewife, student, retired, not working.

cAll baseline symptom scales were standardized to a mean of 0 and variance of 1 in the total sample.

dConsistently lowest value across samples.

eAll week 1 and week 3 symptom scores were standardized to the week 0 means and standard deviations.

fChange scores for continuous variables are in the metric of the standardized week 0 scores.

gThese 3 categories are nested. So the “higher or the same” category includes all those either higher or the same and the “slightly lower or more” includes all those higher as well as those the same and those with negative change scores no less than 0.25 standard deviations (of the change score) below the mean.

hThe dummy variables were selected based on inspection of the week 1, week 3, and weeks 1-3 change distributions of the cross-classification of frequency with severity of side effects.

**References**

Chen, T., & Guestrin, C. (2016). XGBoost: a scalable tree boosting system. *arXiv* [Preprint]. Retrieved from https://arxiv.org/abs/1603.02754

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, *38*(4), 367-378. doi:10.1016/S0167-9473(01)00065-2

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of* *Statistical* *Software,* *33*(1), 1-22. doi:10.18637/jss.v033.i01

Gelman, A., Jakulin, A., & Pittau, M. G. (2008). A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics*, *2*(4), 1360-1383. doi:10.1214/08-AOAS191

Hastie, T., & Tibshirani, R. (1986). Generalized additive models. *Statistical* *Science*, *1*(3), 297-318. doi:10.1214/ss/1177013609

Kooperberg, C. (2015). polspline: polynomial spline routines. R package version 1.1.12. Retreived from https://cran.r-project.org/web/packages/polspline/polspline.pdf

Luedtke, A. R., & van der Laan, M. J. (2017). Evaluating the impact of treating the optimal subgroup. *Statistical Methods in Medical Research*, *26*(4), 1630-1640. doi:10.1177/0962280217708664

Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A (General),* *135*(3), 370-384. doi:10.2307/2344614

Scholkopf, B., Platt, J. C., & Shawe-Taylor, J. (2001). Estimating the support of a high-dimensional distribution. *Neural* *Computation*, *13*(7), 1443-1471. doi:10.1162/089976601750264965

van der Laan, M. J., Polley, E. C., & Hubbard, A. E. (2007). Super learner. *Statistical Applications in Genetics and Molecular Biology*, *6*, Article25. doi:10.2202/1544-6115.1309

Wright, M. N., & Ziegler, A. (2017). ranger: a fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software,* *77*(1), 1-17. doi:10.18637/jss.v077.i01