Factors Associated with Medicine Timing Effects: A Meta-analysis

Marc D. Ruben¹, Lauren J. Francey¹, Gang Wu¹, David F. Smith²,³, Garret A. FitzGerald⁴,⁵,⁶, Jacob J. Hughey⁷,⁸, John B. Hogenesch¹

¹ Divisions of Human Genetics and Immunobiology, Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, USA.
² Divisions of Pediatric Otolaryngology and Pulmonary and Sleep Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, USA.
³ Department of Otolaryngology-Head and Neck Surgery, University of Cincinnati School of Medicine, Cincinnati, USA.
⁴ Department of Systems Pharmacology and Translational Therapeutics, at the University of Pennsylvania Perelman School of Medicine, Philadelphia, USA
⁵ Department of Medicine, at the University of Pennsylvania Perelman School of Medicine, Philadelphia, USA
⁶ Institute for Translational Medicine and Therapeutics (ITMAT), at the University of Pennsylvania Perelman School of Medicine, Philadelphia, USA
⁷ Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, USA.
⁸ Department of Biological Sciences, Vanderbilt University Medical Center, Nashville, USA.

Abstract

Importance Clinical evidence suggests that the time of day of treatment can affect outcomes in many different diseases, but this information is dispersed, imprecise, and heterogeneous. Consequently, practice guidelines and clinical care recommendations seldom specify intervention time.

Objective To understand the sources of variability and summarize clinical findings on the time of day effects of medicine.

Data Sources A systematic search of Pubmed, Google Scholar, and ClinicalTrials.gov for “chronotherapy” OR “time of administration”.

Study Selection Any clinical study since 2000, randomized or observational, that compared the effects of treatment at different times of day. We included pharmacologic or surgical interventions having at least one continuous outcome.

Data Extraction and Synthesis For selected studies, we extracted the mean and variance of each time-of-day treatment group. From these, we computed the standardized mean difference (SMD) as the measure of timing effect. Where a study reported multiple outcomes, we selected a single outcome based on a defined order of priority.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
**Main Outcomes and Measures**  We estimated overall pooled effect size and heterogeneity by a random effects model, followed by outlier detection and subgroup analyses to evaluate how study factors, including drug, design, outcome, and source, associate with timing effect.

**Results**  78 studies met the inclusion criteria, comprising 48 distinct interventions over many therapeutic areas. We found an overall effect of time on clinical outcomes but with substantial heterogeneity between studies. Predicted effects range from none to large depending on the study context. Study size, registration status, and source are associated with the magnitude of effect. Larger trials and those that were pre-registered have markedly smaller effects, suggesting that the published record overstates the effects of the timing of medicine on clinical outcomes. In particular, the notion that antihypertensives are more effective if taken at bedtime draws disproportionately from one source in the field, which consistently detects larger effects than the community average. Lastly, among the most highly studied drug timing relationships, aspirin’s anti-clotting effect stands out, consistently favoring evening over morning dosing.

**Conclusions and Relevance**  While accounts of drug timing effects have focused on yes/no, appreciating the range of probable effects may help clarify where ‘circadian medicine’ meets the threshold for clinical benefit.

**Introduction**

Half of human genes and the physiology they control are expressed with a 24 h rhythm with implications for medicine \(^1\,^2\). Treatment time effects can in theory arise from circadian influences on pharmacokinetics, pharmacodynamics, or underlying physiology. Indeed, dozens of clinical trials suggest that time of dosing can affect responses in cancer, heart, lung, and other diseases. Even vaccines may be influenced by the timing of their administration \(^3\,^4\). Yet, outside of sleep medicine, time of day is rarely considered in medical practice. The most recent WHO Model List of Essential Medicines makes no mention of dosing time. In view of notable advances in circadian biology research, why are there so few examples of clinical use?

Dosing time has been most studied in hypertension, and randomized trials seem to converge on a few fundamentals \(^5\,^6\). For example, multiple randomized trials show that short-acting antihypertensives taken at bedtime better restore the normal overnight drop in blood pressure compared to morning dosing \(^6\). But even here there is uncertainty. Trials differ by drug, design, and endpoint, and not all findings point in the same direction. Of the dozens of antihypertensive medications, none have garnered an FDA time-of-day directive, and the importance of dosing time is disputed \(^8\,^9\). In fact, we have only a superficial understanding of how much treatment time affects cardiac outcomes; this question applies to all indications. A history of trials report treatment time differences, but information is heterogeneous and imprecise. There is a need to synthesize the evidence.

To understand the sources of variability and more adequately summarize clinical findings on the time-of-day uses of medicine, we analyzed 78 studies that compared treatment at different times
of day. This comprised 48 distinct interventions over many therapeutic areas. To our knowledge, this is the largest such analysis to date.

**Methods**

**Study Identification**

We conducted a systematic search of PubMed, ClinicalTrials.gov, and Google Scholar for clinical studies that compared the effects of treatment administered at different times of day in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline. The first search was performed in March 2020. We used the search terms “chronotherapy” or “time of administration” and included only articles in English.

After title and abstract screening, the full texts of potentially relevant studies were reviewed and selected if they evaluated the effects of the same treatment given at different times of day. We allowed any pharmacologic or surgical intervention as long as the study tested at least one continuous outcome measure. Because the aim of our study was a broad evaluation of medicine timing effects, we allowed any study design—randomized or observational, and any disease context, study population, and outcome/endpoint—as long as the effect mean and variance of each time-of-day group were available. The only exception to this were studies addressing sleep disorders, which we excluded.

We screened the hits from the initial search for relevant references. We performed a final search of the same three databases in June 2021 (Fig. 1 Supplement). Our final selection includes studies meeting the above criteria published since January 1st 2000.

**Data Extraction and Methodological Quality Assessment**

This study was not submitted for institutional review board approval because it did not involve individual patient information and all data extraction was from publicly available sources. We collected study characteristics: drug ingredient, drug class, drug half-life, study design, registration status, treatment setting, patient demographics, investigator source, and primary and secondary outcomes. For study outcomes, we extracted the size, mean, and variance of each time-of-day treatment group.

While we did not grade the quality of each study a priori, we collected study characteristics used by established methods of grading. We sought to understand sources of heterogeneity in the body of literature and our analyses consider several commonly used proxies for quality.

**Statistical Analysis**

We computed Cohen’s d standardized mean difference (SMD) for each continuous outcome, from each study. Specifically, we subtracted the mean post-evening-treatment outcome from the mean post-morning-treatment outcome and divided the result by the pooled standard
deviation for both groups (Eq. 1 and Eq. 2). Where more than two times of day were tested, we computed the SMD between the two groups with the largest raw mean difference.

\[
SMD_{\text{between times}} = \frac{\bar{x}_{\text{evening}} - \bar{x}_{\text{morning}}}{s_{\text{pooled}}} 
\]

Eq. 1 standardized mean difference

\[
s_{\text{pooled}} = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1-1) + (n_2-1)}}
\]

Eq. 2 pooled standard deviation

\[n_1, s_1: \text{sample size and standard deviation of evening group}\]
\[n_2, s_2: \text{sample size and standard deviation of morning group}\]

To correct for small-sample bias in Cohen’s SMD, we converted Cohen’s SMD to Hedge’s SMD using the R package ‘esc’. Standard error (SEM) of SMD was computed according to Eq. 3. Datafile 1 in the supplement contains the fully curated and coded dataset.

\[
SEM_{SMD_{\text{between times}}} = \sqrt{\frac{n_1 + n_2}{n_1n_2} + \frac{SMD_{\text{between times}}^2}{2(n_1 + n_2)}}
\]

Eq. 3 SMD standard error

We prioritized end-of-study means as opposed to baseline subtracted, ie. change score, means. In rare cases where means were baseline-covariate-adjusted, we prioritized these, as reported in Data file 1 in the supplement. In cases where within-group variances were expressed as standard error of the mean (SEM) or 95% confidence intervals, we computed SD according to methods described in the Cochrane Handbook for Systematic Reviews of Interventions. If only median and interquartile range were reported, mean and SD were approximated.

Where a study reported multiple continuous outcomes, we selected a single outcome based on the following order of priority: (i) significant effect \((p < 0.05)\) and a designated primary endpoint, (ii) designated primary endpoint, (iii) significant effect, and lastly (iv) none of the above. If multiple outcomes remained, we selected the one with the largest SMD to include in the meta-analysis.

In a small number of cases, we excluded certain outcomes from eligibility (Data file 1 in the supplement). For example, whereas awake/asleep ambulatory blood pressure are well accepted clinical endpoints in a study of hypertension, mean blood pressure over an arbitrarily defined window, eg. 11AM to 3PM, is not.

In our primary analysis, we estimated the overall SMD across studies by pooling the individual study SMDs using a random-effects model with a DerSimonian-Laird estimator via the R package ‘meta’. To identify potential sources of heterogeneity between studies, we
recalculated the meta-analysis results N times, each time leaving out one study, and then
prespecified subgroup analyses were performed using a mixed-effects model via the R package
'dmetar' \(^{23}\). The mixed-effects model estimates the variability within subgroups as a random
effect and the variability between subgroups as a fixed effect. A p-value for difference between
groups was calculated via Q-test. Subgroups evaluated include study design, registration status,
source, and drug kinetics. Drug half-lives were extracted from DrugBank 5.0 \(^{24}\), FDA-label, or
research publication. We evaluated for small-study bias via funnel plot \(^{25,26}\).

Our study includes two secondary analyses. The first compares raw mean differences among a
set of blood pressure studies with outcomes on the same unit scale, by drug type and kinetics,
patient demographics, and the study source. The second analysis evaluated the reproducibility
of interventions tested in more than one study.

**Results**

Our search revealed 172 potentially relevant reports, of which 78 fulfilled our eligibility criteria
(Fig. 1 Supplement). The 78 studies encompass 48 distinct treatments tested at different times
of day in multiple disease settings (Figure 1A, Data file 1 Supplement, Methods). Most studies
tested only two time points, typically morning and evening.

**Standardized effects of time across interventions**

We hoped to gain a sense of which interventions were most susceptible to time of day effects.
We computed the standardized mean difference (SMD) \(^{14,17}\) for each study (Fig. 1B). SMD
provides a measure of effect size that is relative to variance within the study. Specifically, we
subtract the mean of one time-of-day group from the other and divide the result by the pooled
standard deviation (SD) (Methods). Effect sizes are given as a fraction of the standard deviation.
This measure is independent of the units or scale in the original study and allows for
comparisons across studies. A study with SMD = 0.5 indicates the two groups differ by 0.5 SD,
SMD = 1 indicates the two groups differ by 1 SD, and so on. Where a study reported multiple
outcomes, we selected a single outcome based on an order of priority described in Methods.
SMDs range from -0.16 to 6.7 (Fig. 1B), with positive values indicating that evening dosing was
favorable to morning dosing, and vice versa. Evening dosing was favorable in most studies
(68%). SMD alone, however, does not indicate sampling error, as studies differ in precision. To
estimate the overall effect of time, we pooled all studies together under a random effects model
which weights studies based on their estimated precision (Fig. 1C). The pooled estimate is
therefore a measure of the mean of a distribution of true effects \(^{17,27}\). The estimated absolute
SMD of 0.41 (95%CI: 0.34 - 0.49, p < 0.0001) indicates an overall positive effect of time.
However, model indicators \(^{28}\) suggest substantial between-study heterogeneity (Fig. 1C; \(I^2 =
82\%\), ie., the percent of variability not caused by sampling error). The wide prediction interval
\(^{29}\) from -0.12 to 0.94 means that the pooled estimate of 0.41 is unlikely to be robust in every
context. We expect some future studies will find little or no timing effect, whereas others will find
large effects. This led us to ask which studies were more or less influential in this prediction.
Leave-one-out analysis identifies an investigator source of heterogeneity

To identify sources of heterogeneity between studies we recalculated the meta-analysis results N times, each time leaving out one study. This allows us to detect the contribution of each study to the overall pooled effect (Fig. 2A, left). Specifically, the ‘difference in fit(s)’ (DFFITS) metric quantifies how much and in which direction the pooled effect changes when a study is omitted. Among the 12 most effect-lowering studies, nine of them relate to blood pressure, seven of which draw from sites in Northwest Spain (N.W. Spain) involving the same investigator (Fig. 2A, middle). This source is responsible for 13 of the 39 blood pressure studies in this meta-analysis, with the other 26 studies spread across different investigative groups and locations (Data file 1 Supplement). Using a mixed-model approach that pools the effect of subgroups and compares between them, the pooled effect from N.W. Spain is 0.80 (95%CI: 0.52 - 1.1) compared to 0.33 (95%CI: 0.24 - 0.41) for all others (p < 0.005) (Fig. 2B). We explore this further in a later section.

Timing effects are associated with registration status

We evaluated whether there are other study contexts where effects are lower or higher. Using a mixed model approach as above, we tested the impact of study pre-registration and design. Clinical study registration provides transparency and discourages publication bias. Because only 34 of 78 studies in this analysis were prospectively registered on ClinicalTrials.gov or an international equivalent, we speculated that the overall evidence base may be distorted. The idea is that without a public record, negative or null trial results may be less likely to be published. Indeed, the pooled effect of pre-registered studies is 0.26 (95%CI: 0.19 - 0.34) compared to 0.57 (95%CI: 0.43 - 0.72) for non-registered studies (p < 0.001) (Fig. 2C). The pooled effect of pre-registered studies becomes even smaller (0.16, 95%CI: 0.09 - 0.22) if we omit all studies from N.W. Spain, several of which were pre-registered.

We found no evidence of association between study design and effect (Fig. 1A Supplement). Although parallel trials had a larger effect than crossover or retrospective designs, this difference can be explained by N.W. Spain studies, most of which were parallel designs.

A small-study effect may indicate publication bias

There is a tendency for smaller studies in a meta-analysis to show larger effects. One possible cause is reporting bias, whereby small negative studies are less likely to be published. Funnel plots are commonly used to investigate small-study bias. The principle is that as studies increase in size and precision, their location around the pooled average effect narrows, forming a funnel. In the absence of bias, small (less precise) study effects should scatter widely, with spread narrowing for larger (more precise) studies. Our analysis detects asymmetry indicative of small-study bias (Fig. 2D). Specifically, the largest studies fall to the left of the pooled effect. This imbalance, together with the prior observation that pre-registered studies have smaller effect sizes, suggests that the published record overstates medicine timing effects.
The evidence for medicine timing in hypertension

Half of the studies in this meta-analysis concern the timing of antihypertensives. More than one-third of these antihypertensive studies originate from N.W. Spain. Understanding why findings from N.W. Spain appear as outliers is important from both a research and clinical perspective.

We first evaluated whether N.W. Spain’s large effects may be due to lower study variances. SMD is scale-free and not influenced by the units of measurement. However, choices made by an investigator, such as enrolling a narrow patient population, can reduce study variance, which may cause SMD to appear larger. Therefore, we computed raw mean differences (RMD = mean mmHg_{morning} – mean mmHg_{evening}) for each of the 38 studies with blood pressure outcomes on the same scale. We focused on nighttime blood pressure because of its prognostic and circadian relevance. Overall, RMDs suggest that evening dosing is preferable to morning dosing for lowering nighttime blood pressure (Fig. 3A). However, the median RMD from N.W. Spain (5.5 mmHg) is markedly larger than other studies (1.7 mmHg). Put another way, N.W. Spain is also an outlier by RMD, as it was by SMD. Perhaps interestingly, N.W. Spain studies do have lower variances, ie. pooled standard deviations (Methods) (Fig. 3A) on average, but this is unlikely to explain the large discrepancy in RMD.

Can differences in the types of drugs studied explain the discrepancy? N.W. Spain has larger RMDs for all six mono- or dual-therapies that were also studied by at least one outside group (Fig. 1B Supplement). Further, the distributions of study drug half-lives are virtually indistinguishable between N.W. Spain and others (Fig. 3B, top left). Therefore, differences in the nature of the drugs studied cannot explain the discrepancy.

We next evaluated patient composition. N.W. Spain consistently enrolled a larger share of younger and female patients than other studies (Fig. 3B, top middle and right panels). This difference in study makeup is dramatic, and sex-specific effects of drug timing have been observed in other contexts. However, closer inspection of the data refutes this idea, as neither sex nor age track with RMD (Fig. 3A, bottom middle and right panels). In sum, the idea that antihypertensives, even short-acting, are more effective if taken at night is only weakly supported by the broader research community.

Aspirin’s time-dependent anti-coagulative effect stands out

This meta-analysis includes studies in many conditions other than hypertension. We evaluated if there are particular areas of medicine with orthogonal lines of support for the importance of time. There were too few studies per outcome to justify regression-based subgroup analyses, so we simply visualized effect (SMD) versus precision (SEM) for each of the 38 non-hypertension studies. With the exception of airway patency, where studies can be characterized as high precision but small effect, there are no remarkable differences between outcomes (Fig. 4A, left).

We next assessed reproducibility, focusing on the five intervention::outcome pairs tested in more than one study: (i) aspirin::coagulation, (ii) levothyroxine::thyroid function, (iii) insulin
glargine::glucose levels, (iv) simvastatin::lipid levels, and (v) atorvastatin::lipid levels (Fig. 4A, right). Simvastatin and aspirin were consistently more effective when taken in the evening, whereas insulin glargine was more effective when taken in the morning. Levothyroxine and atorvastatin were inconsistent, favoring evening dosing in some studies and morning in others.

All aspirin studies in this meta-analysis tested low 75-100 mg 'cardiopreventive' doses. Aspirin has well established anti-thrombotic efficacy and low-dose aspirin is standard-of-care for secondary prevention of cardiovascular disease \(^{47}\), taken daily by millions of patients. The understanding that its effect is sensitive to timing is not reflected in current dosing guidelines. Given the potential clinical importance, we examined all available evidence. As a reminder, we selected only a single outcome per study to include in the meta-analysis, based on a defined order priority (Methods). However, secondary and other endpoints in a study are also informative. To scrutinize all available evidence, we compared all of the reported outcomes of aspirin in coagulation for all aspirin studies in this meta-analysis (Fig. 4B). Eight out of ten outcomes favored evening over morning dosing, across different assays of platelet inhibition.

Our prior work suggested that short-acting drugs are more likely to show timing effects \(^{7}\). This assessment relied on a binary yes/no measure of effect, ie. significance \(p< 0.05\). We re-evaluated the impact of half-life via meta-regression \(^{22}\) onto SMD, and found half-life to be a surprisingly poor predictor of effect size (slope = -0.0008, \(p = 0.8\), \(R^2 = 0\)) (Fig. 2C Supplement). Of note, our prior assessment included several studies of short-acting anti-cancer agents that were not part of this study because they reported only binary outcomes.

**Discussion**

Treatment time effects can in theory arise from time-of-day influences on pharmacokinetics, pharmacodynamics, or underlying physiology. As genome-wide circadian datasets grow, so have predictions of medicine timing effects. In the past five years, ~150 reviews mentioned 'chronotherapy' or 'circadian medicine' in the title or abstract, outpacing clinical evidence. In this meta-analysis, we estimate a summarized effect of time of 0.41 standard deviations. What does this mean? While there is no straightforward way to assign reference values, an analysis of the most common drugs versus placebo found effects from SMD=0.1 to SMD=1.4 \(^{48}\). In this context, timing effects from several studies were as large as many drug-versus-nothing effects. Unfortunately, the evidence for many of the interventions draws from a single trial of low precision. Which interventions are well-supported clinical opportunities that deserve focus?

Evidence from aspirin studies in this meta-analysis warrants future randomized controlled trials of aspirin timing in the prevention of cardiac events. Aspirin inhibits platelet reactivity and is widely used for cardiovascular disease prevention for this reason \(^{49}\). Aspirin’s efficacy is attributed to its inhibition of cyclo-oxygenase-1 (COX-1) leading to inhibition of thromboxane A2 production by platelets \(^{50}\). Interestingly, because the COX-1 inhibition is irreversible, aspirin’s therapeutic effect is traditionally thought to be sustained throughout the dosing interval once steady state is achieved. What, then, accounts for the dosing time effect on platelet reactivity
observed in trials in this meta-analysis? One possibility is drug kinetics and platelet turnover. Aspirin has a short ~1 h half-life and inhibits only the platelets present at the dosing time. New platelets are released at a rate of 10% per day\textsuperscript{51,52} and it has been shown that the antiplatelet effect of aspirin declines during the 24 h dosing interval\textsuperscript{53,54}. Alternatively, the variable inhibition of other coagulation factors, such as prostaglandin I2, may modulate aspirin’s therapeutic platelet inhibition. Importantly, it is well established that platelet reactivity and coagulation peak in the morning\textsuperscript{41,42}. Approximately 300K recurrent cardiovascular events occur in the United States every year, with an excess of 40% during the morning hours\textsuperscript{55}. Even a small reduction in the morning peak could translate to a large decrease in recurrent events per year\textsuperscript{56}.

Studies of aspirin in platelet inhibition\textsuperscript{57} underscore another point: the better time for treatment should be interpreted in the context of what was measured. For example, whereas morning dosing of proton pump inhibitors (PPIs) was more effective for \textit{activity}-induced gastroesophageal reflux, evening dosing was more effective for \textit{nocturnal} reflux\textsuperscript{58}. Constriction of the airway and pain also tend to peak at night. In this sense, optimal timing may be a matter of aligning drug exposure with a 24 h rhythmic symptom.

No area of drug timing has more published studies and debate than hypertension. Daily rhythms in blood pressure are well-documented. Disruption of these rhythms, in particular the loss of normal nighttime ‘dipping’, is associated with increased cardiovascular risk\textsuperscript{38–40}. Evidence that bedtime ingestion of antihypertensives is more effective for controlling nighttime blood pressure is therefore of wide interest. And interestingly, the greatest effects are seen with aspirin, findings that have been reproduced in mice\textsuperscript{59}. However, the clinical evidence draws largely from one source with effect sizes that are difficult to reconcile with the broader field. While confusing, we expect that continued advances in understanding the time-of-day-dependent mechanisms of cardiovascular function will clarify the role of time in treatment of heart disease.

\textbf{Limitations}

SMD is an imperfect metric. Although insensitive to measurement units and thus ‘scale-free’, factors that influence the variance in a particular study can impact SMD and thereby compromise between-study comparisons. In addition, SMD is not robust to study outliers. Because outliers more strongly affect the variance than the mean, the presence of large outliers in a study can make SMD appear smaller than an analogous study without outliers. Alternative metrics that may better handle these issues require patient-level data, which was not available.

Where a study reported multiple outcomes, we selected a single one to use following an order of priority that favored statistical significance, as described in Methods. Our approach may therefore present a best-case scenario for timing effects. Alternative approaches that integrate \textit{all} reported effects from a single study may be desirable but are challenging without patient-level data. Additionally, this meta-analysis did not consider binary outcomes, such as incidence of toxicity or survival, excluding ~25 studies of chemotherapy timing in cancer. We collected data from these studies and provided it as a resource (Data file 1 in the Supplement).
Conclusions

Despite many reports that attention to time-of-day can improve therapeutic outcomes, there is a critical need for both rigorous randomized controlled trials and analysis of real-world health data. Mechanistic insights from chronobiology will continue to drive hypotheses, but the burden of proof is clinical benefit. Published accounts of timing effects have tended to focus on yes/no significance. Future efforts to understand the range of likely effects should help clarify instances where consideration of time of day has the greatest medical potential.

Main Figures
Fig. 1. Effects of intervention time in clinical studies. (A) 78 studies since 2000 (black triangles) that tested the impact of time of day on treatment outcomes. Data file 1 in the Supplement holds the full curated dataset. (B) Standardized mean difference (SMD) grouped by outcome measure.

random effects model

model estimate = 0.4 [0.33–0.47]

prediction interval [−0.12–0.92]
the type of outcome measure, where \( \text{SMD} = \bar{x}_{\text{evening}} - \bar{x}_{\text{morning}} \) divided by the pooled standard deviation. *Filled circles*: outcomes that were reported in the original research as significant at \( p < 0.05 \) by null hypothesis testing. SMD values were capped at 1.5 for visualization, impacting two blood pressure studies with SMDs of 6.7 and 2.9, both set to 1.5 in this plot. All raw data is available in Data file 1 in the Supplement. *(C)* Absolute SMD with 95% confidence interval. *Blue diamond*: pooled effect under a random effects model. *Red bar*: Prediction interval within which future studies can be expected to fall. *Grey vertical line*: zero timing effect. *Inset table*: between-study heterogeneity. Note: We removed one extreme outlier study for visualization purposes only, Farah et al 2013, with SMD = ~7.
**Fig. 2. Study source and pre-registration condition effect size.** (A) Size and direction of the change in the pooled effect when each study is omitted. DFFITS on the y-axis is expressed in standard deviations; larger values indicate decreased pooled effect when the study is omitted, and vice versa. Top 10 increases and decreases labeled. DFFITS values were capped at 0.35 for visualization, impacting two blood pressure studies with values of 6.7 and 2.9. (B) Subgroup analysis of Northwest Spain (N.W. Spain) vs. other studies. Blue line: pooled effect of all studies. (C) Subgroup analysis of pre-registered versus non-pre-registered for all studies (left panel) or omitting N.W. Spain (right panel). Blue line: pooled effect for all studies. (D) Study precision (SEM_{SMD}) versus effect (SMD). Blue vertical line: pooled effect. Blue diagonal lines: pseudo 95th percentile confidence intervals around the pooled average as precision decreases from zero. P-value from Egger’s test for asymmetry of the funnel plot. Absolute SMD values were capped at 2.5 for visualization only, impacting two blood pressure studies, SMD = 6.7 and 2.9, both set to 2.5 in this plot. **Drug class abbreviations:** ARB: angiotensin receptor blocker, CCB: Ca^{2+} channel blocker, nAChRi: anticholinergic, NSAID: non-steroidal anti-inflammatory, mixed AH: combination of antihypertensives, ACEi: angiotensin converting enzyme inhibitor; A2-agonist: alpha-2 adrenergic receptor agonist, CORT: corticosteroid, antihistH1: antihistamine, GlucagonR-i: glucagon-like peptide-1 receptor agonist, VKA: vitamin K antagonist.
Fig. 3. Factors influencing dosing time effects on nighttime blood pressure. (A) 38 studies that tested the effect of morning versus evening dosing on nighttime systolic blood pressure. Effect is shown on the vertical axis as raw mean difference (RMD) where $RMD = \text{mean mmHg}_{\text{morning}} - \text{mean mmHg}_{\text{evening}}$. Study variance is shown on the horizontal axis as pooled standard deviation (Methods). Marginal boxplots summarise how RMD and variance are distributed. (B) Top row: density plots indicate relative numbers of studies at different values of drug half-life, mean patient age, and percent sex. Orange lines: N.W. Spain; grey lines: all other studies. Bottom row: regressing RMD onto drug half-life, mean patient age, and percent sex. Smooth lines fit by the generalized additive model $y \sim s(x, \text{bs} = "cs", \text{knots} = 5)$, where $y = \text{RMD}$, $x = \text{covariate}$, $\text{cs} = \text{cubic spline basis function}$. Grey shading: 95% confidence intervals. Drug half-lives obtained from DrugBank 5.0, FDA-label, or research publication. For studies with drug combinations, the shortest-acting ingredient is considered.
Fig. 4. Assessment of reproducibility. (A) Left: absolute SMD versus precision (SEM) for all meta-analysis studies, excluding blood pressure outcomes (n = 38). Point shape: whether morning or evening treatment was more effective. Right: the five intervention::outcome pairs tested in more than one study: (i) aspirin::coagulation, (ii) levothyroxine::thyroid function, (iii) insulin glargine::glucose levels, (iv) simvastatin::lipid levels, and (v) atorvastatin::lipid levels. Pairs are encircled to highlight group coherence. (B) All outcomes of aspirin in coagulation for each of the five studies of aspirin in this meta-analysis (n = 10). Data presented here deviate from our meta-analysis inclusion criteria of one outcome per study (described in text, Methods). Point shape: study reference. Abbreviations: ARU: aspirin relative units from VerifyNow test.
Supplement

Selection of studies included in meta-analysis

Studies found in database searches
- PubMed (1,617)
- Google Scholar (8,840)
- ClinicalTrials.gov (23)

10,308 studies did not measure the effects of the same treatment given at different times of day. Reviews, editorials, commentaries, etc. were excluded

172 studies downloaded

94 studies excluded:
- no continuous outcome
- no effect mean and/or variance
- suspected overlap of subjects with another included trial
- prior to year 2000

78 studies selected

Fig 1 Supplement. Selection of studies included in the meta-analysis (Methods).
Fig 2 Supplement. Influence of study features on timing effects (A) Mixed-effects model pools the effect of each study design subgroup and compares between them. (B) Raw mean difference (RMD) for each distinct drug or drug combination that tested the effect of morning vs. evening dosing on nighttime systolic blood pressure. (C) Meta-regression of SMD onto drug half-life. **Point size:** study weight, ie., the contribution of each study to the estimated pooled effect. **Model coefficients:** R² = percent of between-study variation explained by the model, I² = residual heterogeneity, or the remaining variability between-studies not captured by half-life predictor, and p-value for difference between groups via Q-test. **Drug class abbreviations:** mACHRi: anticholinergic, A2-agonist: alpha 2 adrenergic receptor agonist, CORT: corticosteroid, GlucagonR-i: glucagon-like peptide-1 receptor agonist, VKA: vitamin K antagonist, SERM: selective estrogen receptor modulator; MRA: mineralocorticoid receptor antagonist.
# 78 Meta-analysis Studies

<table>
<thead>
<tr>
<th>OUTCOME AREA</th>
<th>REF</th>
<th>LINK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood pressure</td>
<td>Krasinska 2020</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Hermida 2019</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Zappe 2015</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>Ruben 2019</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Rahman 2013</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>Tiwari 2020</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Bonten 2015</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Wang 2013</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Ayala 2013</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Zeng 2011</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Kasiakogias 2015</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Asmar 2011</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>Hermida 2010</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Hoshino 2010</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Hermida 2010</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Hermida 2009</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Hermida 2009</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Hermida 2009</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>H. Mori 2013</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Ushijima 2015</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Hermida 2009</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Hermida 2007</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Hermida 2005</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>Hermida 2005</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Qiu 2003</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Serinel 2017</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>Hjortkjær 2016</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>Kitahara 2004</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Hermida 2008</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Hermida 2005</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Kohno 2000</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Farah 2013</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Rossen 2014</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>Fujiwara 2017</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>Huangfu 2015</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>Kario 2016</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>Dimitrov 2012</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Suomela 2015</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>Khodadoustan 2017</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>Coagulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krasinska 2018</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>Racca 2019</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>van Diemen 2021</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Brunner-Ziegler 2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garrison 2020</td>
<td>101</td>
<td></td>
</tr>
<tr>
<td>NCT00515021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ando 2013</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>Lipid levels</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chipchura 2018</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td>Ozaydin 2006</td>
<td>104</td>
<td></td>
</tr>
<tr>
<td>Tharavanij 2010</td>
<td>105</td>
<td></td>
</tr>
<tr>
<td>Lafeber 2015</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>Wallace 2003</td>
<td>107</td>
<td></td>
</tr>
<tr>
<td>Lund TM 2002</td>
<td>108</td>
<td></td>
</tr>
<tr>
<td>Kim 2013</td>
<td>109</td>
<td></td>
</tr>
<tr>
<td>Yoon 2011</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>Plakogiannis 2005</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>Poulter 2018</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td>Thyroid function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bach-Huynh 2009</td>
<td>113</td>
<td></td>
</tr>
<tr>
<td>Bolk 2010</td>
<td>114</td>
<td></td>
</tr>
<tr>
<td>Elliott 2001</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>Rajput 2011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Packard 2020</td>
<td>116</td>
<td></td>
</tr>
<tr>
<td>Buttgerfeit 2008</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>Airway patency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calverley 2003</td>
<td>118</td>
<td></td>
</tr>
<tr>
<td>Zetterstrom 2008</td>
<td>119</td>
<td></td>
</tr>
</tbody>
</table>
Table 1 Supplement. All studies included in this meta-analysis.

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCT03108027</td>
<td>2016</td>
</tr>
<tr>
<td>Kempsford 2016</td>
<td></td>
</tr>
<tr>
<td>Kempsford 2013</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>Ren 2019</td>
<td>2019</td>
</tr>
<tr>
<td>de Bree 2020</td>
<td>2020</td>
</tr>
<tr>
<td>Montaigne 2017</td>
<td></td>
</tr>
<tr>
<td>Long 2016</td>
<td>2016</td>
</tr>
<tr>
<td>Ahren 2013</td>
<td>2013</td>
</tr>
<tr>
<td>Haye 2005</td>
<td></td>
</tr>
<tr>
<td>Patti 2020</td>
<td></td>
</tr>
<tr>
<td>Cheeseman 2007</td>
<td>2007</td>
</tr>
<tr>
<td>Lucidi 2015</td>
<td>2015</td>
</tr>
<tr>
<td>Garg 2004</td>
<td>2004</td>
</tr>
<tr>
<td>Newcorn 2013</td>
<td>2013</td>
</tr>
</tbody>
</table>

**Article Information**

**Data deposition:**
Data and code for reproducing the analyses in this study have been deposited in GitHub (https://github.com/MarcDRuben/MetaAnalysis).

**Corresponding Authors:**
Marc D. Ruben PhD. Divisions of Human Genetics and Immunobiology, Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, USA (mruben1975@gmail.com).

John B. Hogenesch, PhD. Divisions of Human Genetics and Immunobiology, Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, USA (hogenesch@gmail.com).

**Author Contributions:** Dr. Ruben had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Concept and design:* Ruben

*Acquisition, analysis, or interpretation of data:* All authors

*Drafting of the manuscript:* Ruben, Francey, Wu, Hughey
Critical revision of the manuscript for important intellectual content: All authors

Statistical analysis: Ruben, Hughey, Wu

Obtained funding: Hogenesch

Supervision: Hogenesch, Hughey

Conflict of Interest Disclosures: None.

Funding/Support: This work is supported by the National Institute of Neurological Disorders and Stroke (2R01NS054794 to JBH and Andrew Liu), the National Heart, Lung, and Blood Institute (R01HL138551 to Eric Bittman and JBH), and National Cancer Institute (1R01CA22748501A1 to Ron Anafi and JBH).

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.
References


8. Lemmer B, Middeke M. Insufficient reply by Hermida et al. to the critical comments to the MAPEC and HYGIA studies. *Chronobiol Int*. Published online September 1, 2020:1.

9. Kreutz R, Kjeldsen SE, Burnier M, Narkiewicz K, Oparil S, Mancia G. Blood pressure medication should not be routinely dosed at bedtime. We must disregard the data from the HYGIA project. *J Hypertens*. Published online August 17, 2020. doi:10.1097/HJH.0000000000002479


30. Belsley DA, Kuh E, Welsch RE. Regression Diagnostics: Identifying Influential Data and


45. Innominato PF, Ballesta A, Huang Q, et al. Sex-dependent least toxic timing of irinotecan
combined with chronomodulated chemotherapy for metastatic colorectal cancer: Randomized multicenter EORTC 05011 trial. *Cancer Medicine*. Published online 2020. doi:10.1002/cam4.3056


60. Krasnińska B, Paluszkiewicz L, Miciak-Ławicka E, et al. The impact of acetylsalicylic acid


77. Mori H, on behalf of the COMPATIBLE Study Group, Yamamoto H, et al. Comparison of effects of angiotensin II receptor blocker on morning home blood pressure and cardiorenal protection between morning administration and evening administration in hypertensive patients: the COMPATIBLE study. *Hypertension Research.* 2013;36(3):202-207. doi:10.1038/hr.2012.142


85. Hjortkjær HØ, Jensen T, Kofoed KF, et al. Nocturnal antihypertensive treatment in patients with type 1 diabetes with autonomic neuropathy and non-dipping: a randomised,


123. Montaigne D, Marechal X, Modine T, et al. Daytime variation of perioperative myocardial


