ADVANCEMENTS IN HEALTH RESEARCH

DOI: 10.4081/ahr.2026.217

SUPPLEMENTARY MATERIAL

METHODS

S1. Model specification and hyperparameter selection

Random Forest and Gradient Boosting hyperparameters were selected using expanding-window rolling temporal validation restricted to the 48-month training period, ensuring that no holdout information contributed to model selection. Four expanding validation blocks were used: training months 1-24 validated on months 25-30; months 1-30 on 31-36; months 1-36 on 37-42; and months 1-42 on 43-48. Within each fold, the outcome was standardised and the time variable was centred and scaled using only the training months from that fold. Candidate hyperparameter grids were as follows: Random Forest, mtry $\in \{1, 2, 3\}$ and terminal node size $\in \{5, 10\}$, with 750 trees during tuning and 1,000 trees in the final fit; Gradient Boosting, interaction depth $\in \{1, 2, 3\}$, shrinkage $\in \{0.01, 0.05\}$, number of trees $\in \{500, 1{,}000, 1{,}500\}$, and minimum terminal node size fixed at 10. The configuration yielding the lowest mean cross-validation RMSE was refitted on the complete training period and applied unchanged to the holdout set. Predictors were country (factor), calendar month (factor), and scaled continuous time. The random seed was fixed at 12345.

S2. Paired country-block bootstrap

Random Forest and Gradient Boosting were compared using a paired country-block bootstrap. Countries, rather than individual country-month observations, were resampled with replacement (10,000 resamples) to preserve the within-country dependence structure of holdout errors. For each resample, differences in R^2, RMSE, MAE, and absolute MBE were recomputed. Percentile confidence intervals and two-sided p values were derived from the resulting bootstrap distributions.

S3. Permutation importance

Holdout permutation importance was estimated by randomly permuting each predictor in the holdout set and recording the resulting increase in holdout RMSE over 200 repetitions. Permutation importance quantifies the fitted model's predictive reliance on a variable; it should not be interpreted as a causal effect or as a partition of explained variance.

S4. Software

Analyses were performed in R version 4.5.2 (R Foundation for Statistical Computing, Vienna, Austria) using the packages tidyverse, readxl, lubridate, sandwich, lmtest, car, nlme, Kendall, forecast, randomForest, gbm, broom and scales.

Table S1. Google Trends extraction protocol. All settings were held constant across the fourteen country queries.

Parameter	Setting used	Comment
Data source	Google Trends (https://trends.google.com/trends/explore)	Publicly accessible; no account required
Date of extraction	20 July 2026	All 14 countries extracted in a single session
Extracted by	MLGL	Single investigator
Search property	Web Search	Image, News, Shopping, and YouTube search not used
Category	All categories (cat = 0)	No algorithmic topic filter imposed
Query type	Search term "ibuprofen"	Not a Google Trends Topic
Geography	One country per query (ISO 3166-1 alpha-2)	Prevents platform-imposed cross-country renormalisation
Country codes (geo)	FR, DE, IT, ES, TR, GB, NL, PL, CZ, BE, NO, SE, HU, AT	France, Germany, Italy, Spain, Turkey, United Kingdom, Netherlands, Poland, Czech Republic, Belgium, Norway, Sweden, Hungary, Austria
Time range	2021-07-01 to 2026-06-30	60 complete calendar months per country
Temporal granularity	Monthly	Partial trailing month excluded to balance the panel
Sub-region / city filter	None	Not available for the Web Search property at country level
Age / device filter	None	Not offered by the platform
Interface language	English (hl = en-GB)	Governs the interface only; does not restrict queries retrieved
Export	Native CSV download	No manual transformation applied
URL template	https://trends.google.com/trends/explore?date=2021-07-01%202026-06-30&geo=XX&q=ibuprofen&cat=0&hl=en-GB	XX = two-letter country code

Table S2. Application of the country-selection criteria. Countries were included when query volume supported a continuous non-zero 60-month series, internet penetration was high, Google was the dominant national search engine, and the panel covered the main European sub-regions.

Country	Sub-region	Population, millions	Internet use, % aged 16–74	Google search share, %	Complete non-zero 60-month series	Included
France	Western	68.4	95	89	Yes	Yes
Germany	Western	83.4	94	90	Yes	Yes
Italy	Southern	58.9	90	94	Yes	Yes
Spain	Southern	48.6	96	96	Yes	Yes
Turkey	South-Eastern	85.4	88.8	60.21	Yes	Yes
United Kingdom	Western	68.3	98	94	Yes	Yes
Netherlands	Western	17.9	100	92	Yes	Yes
Poland	Central–Eastern	36.6	90	95	Yes	Yes
Czech Republic	Central–Eastern	10.9	94	82	Yes	Yes
Belgium	Western	11.8	96	90	Yes	Yes
Norway	Northern	5.6	100	94	Yes	Yes
Sweden	Northern	10.6	98	92	Yes	Yes
Hungary	Central–Eastern	9.6	94	98	Yes	Yes
Austria	Western	9.2	95	90	Yes	Yes

Note. Internet-use estimates refer to the proportion of individuals aged 16-74 using the Internet in 2024 and were obtained from Eurostat-based country statistics; the corresponding value for Türkiye was obtained from the Turkish Statistical Institute (TURKSTAT). Google search-engine market shares refer to 2024 and were derived from StatCounter Global Stats. Values are presented as reported by the source and may therefore differ slightly in numerical precision across countries.

Table S3. Seasonality metrics by country, derived from robust periodic STL decomposition of the 60-month series. Seasonal strength is FS = max(0, 1 - Var(remainder)/Var(seasonal + remainder)); amplitude is the difference between the largest and smallest mean monthly seasonal component. The winter-summer contrast compares December-February with May-July across the four complete seasonal cycles (2022-2025). The final column is the Benjamini–Hochberg-adjusted p value for the joint HAC-adjusted test of calendar-month effects across the fourteen country tests. RSV, relative search volume.

Country	Peak month	Trough month	Seasonal strength (FS)	Amplitude (RSV points)	Mean winter – summer, RSV (SD)	Cycles with winter > summer	Month effect, q
France	Dec	Jun	0.28	17.5	12.8 (13.4)	3 of 4	<0.001
Germany	Dec	Jun	0.40	21.7	10.9 (15.1)	3 of 4	<0.001
Italy	Dec	Jun	0.46	28.4	19.0 (19.8)	3 of 4	<0.001
Spain	Dec	Jun	0.59	18.2	12.8 (10.9)	3 of 4	<0.001
Turkey	Dec	May	0.23	14.9	7.8 (9.7)	3 of 4	<0.001
United Kingdom	Dec	Jul	0.59	18.3	6.2 (6.0)	3 of 4	<0.001
Netherlands	Feb	Jun	0.46	16.2	7.7 (10.4)	3 of 4	<0.001
Poland	Jan	May	0.72	24.8	20.3 (11.1)	4 of 4	<0.001
Czech Republic	Jan	May	0.29	14.2	11.3 (11.6)	3 of 4	<0.001
Belgium	Dec	Jun	0.60	16.7	11.8 (11.3)	3 of 4	<0.001
Norway	Feb	May	0.20	10.5	7.0 (5.0)	4 of 4	<0.001
Sweden	Dec	Apr	0.48	20.0	11.6 (5.6)	4 of 4	<0.001
Hungary	Dec	May	0.50	25.3	14.5 (7.5)	4 of 4	<0.001
Austria	Dec	Jun	0.52	37.3	16.8 (15.7)	3 of 4	<0.001

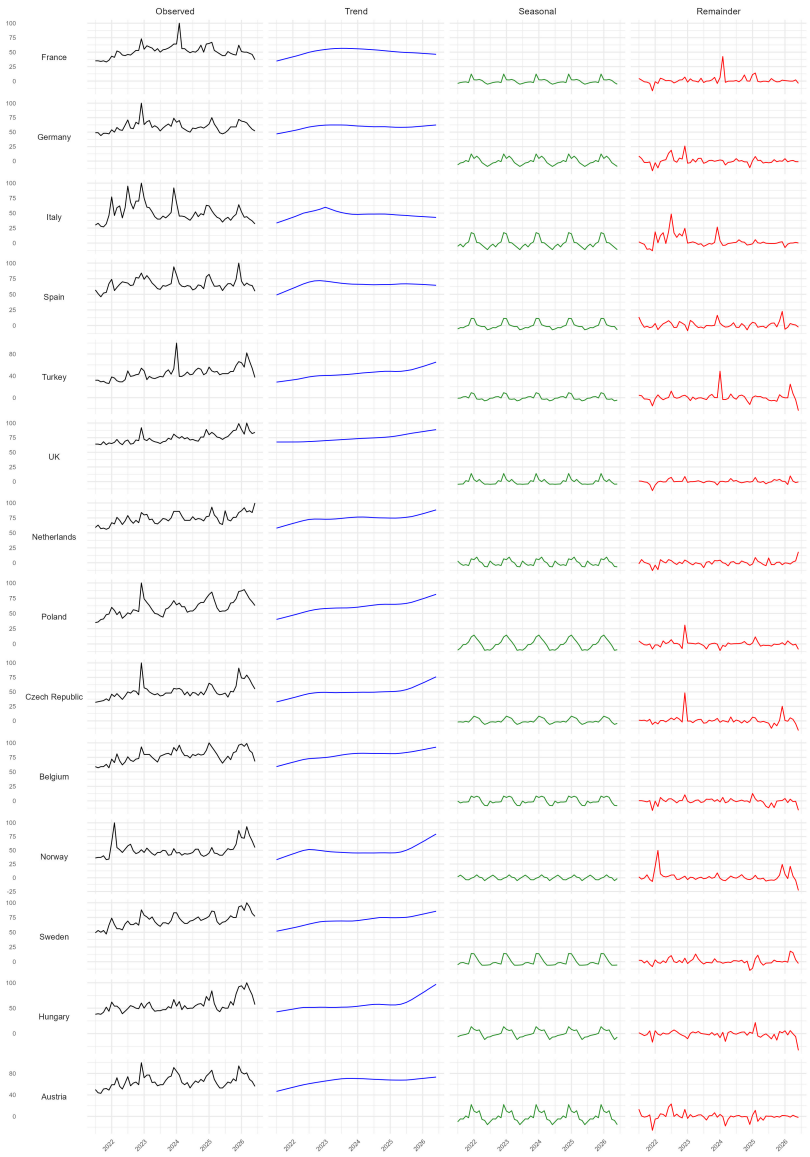
Table S4. Sensitivity analyses for the within-country secular trend. The primary estimate is the calendar-month-adjusted regression with Newey–West HAC standard errors reported in Table 1. q values are Benjamini–Hochberg-adjusted within each method across the fourteen country tests. GLS, generalised least squares with an AR(1) residual structure (maximum likelihood); τ, Kendall's tau; ACF(1), first-order autocorrelation of the primary-model residuals.

Country	Primary HAC slope/year (95% CI); q	Seasonal Mann-Kendall τ; q	GLS AR(1) slope/year (95% CI); q	Residual ACF(1)	Direction / FDR status concordant
France	+1.57 (-2.76 to 5.89); 0.498	+0.05; 0.721	+1.39 (-1.60 to 4.37); 0.414	0.57	Yes / Yes
Germany	+1.00 (-1.94 to 3.94); 0.498	+0.14; 0.294	+1.02 (-1.15 to 3.18); 0.414	0.49	Yes / Yes
Italy	−2.00 (-7.20 to 3.20); 0.498	−0.40; 0.002	−1.10 (-5.63 to 3.44); 0.629	0.63	Yes / No
Spain	+1.02 (-1.49 to 3.54); 0.498	+0.08; 0.601	+0.85 (-1.19 to 2.88); 0.437	0.53	Yes / Yes
Turkey	+5.58 (3.84 to 7.32); <0.001	+0.67; <0.001	+5.46 (3.51 to 7.41); <0.001	0.24	Yes / Yes
United Kingdom	+4.57 (3.48 to 5.66); <0.001	+0.90; <0.001	+4.55 (3.62 to 5.47); <0.001	0.34	Yes / Yes
Netherlands	+4.13 (2.37 to 5.89); <0.001	+0.63; <0.001	+4.84 (3.00 to 6.68); <0.001	0.39	Yes / Yes
Poland	+5.85 (4.11 to 7.59); <0.001	+0.82; <0.001	+5.84 (4.36 to 7.32); <0.001	0.33	Yes / Yes
Czech Republic	+4.56 (1.77 to 7.34); 0.003	+0.55; <0.001	+4.59 (2.37 to 6.82); <0.001	0.39	Yes / Yes
Belgium	+4.66 (2.50 to 6.82); <0.001	+0.64; <0.001	+4.55 (2.90 to 6.19); <0.001	0.43	Yes / Yes
Norway	+2.97 (-1.80 to 7.73); 0.304	+0.25; 0.055	+3.21 (-0.67 to 7.09); 0.144	0.63	Yes / Yes
Sweden	+5.53 (3.68 to 7.38); <0.001	+0.68; <0.001	+5.56 (3.89 to 7.23); <0.001	0.50	Yes / Yes
Hungary	+6.20 (3.18 to 9.22); <0.001	+0.69; <0.001	+6.00 (3.65 to 8.36); <0.001	0.46	Yes / Yes
Austria	+3.07 (0.23 to 5.92); 0.054	+0.38; 0.002	+3.03 (0.90 to 5.16); 0.010	0.44	Yes / No

Note. Direction and FDR status are compared with the primary HAC model. Discordance in FDR status was confined to Austria (primary q = 0.054, significant under both sensitivity models) and Italy (significant monotonic decline under the seasonal Mann–Kendall test only).

Table S5. Cross-country co-movement: the primary analysis (Pearson correlations between STL remainder components) compared with the first-difference sensitivity analysis, for the five strongest and five weakest country pairs of the 91 possible pairs. Across all 91 pairs, the two sets of coefficients were strongly associated (Pearson r = 0.71; Spearman ρ = 0.72), and 51 of 91 remainder correlations remained significant after Benjamini–Hochberg correction. The complete 14 × 14 matrices are available from the corresponding author.

Country pair	r (STL remainders)	Rank of 91	r (first differences)	Rank of 91
Germany - Austria	0.79	1	0.87	1
Poland - Czech Republic	0.76	2	0.81	4
Germany - Poland	0.66	3	0.78	5
Germany - United Kingdom	0.65	4	0.70	10
Germany - Italy	0.61	5	0.62	15
France - Sweden	0.01	87	0.29	70
Germany - Norway	-0.04	88	0.21	81
France - Norway	-0.05	89	0.18	83
Netherlands - Hungary	-0.20	90	0.26	75
Netherlands - Norway	-0.36	91	0.09	89



Supplementary Figure S1. STL decomposition of ibuprofen relative search volume by country, separating the observed series, the long-term trend, the seasonal component, and the remainder. The analysis used robust periodic STL; all components are interpreted within country. This figure was moved from the main text to the Supplementary Material in revision to improve the legibility of the main figure set. RSV, relative search volume; STL, seasonal-trend decomposition by LOESS.