nature food

Analysis

https://doi.org/10.1038/s43016-022-00644-2

Limiting grazing periods combined with proper housing can reduce nutrient losses from dairy systems

Received: 2 June 2022

Accepted: 20 October 2022

R. W. McDowell [1,2] ✉, C. A. Rotz[3], J. Oenema[4] & K. A. Macintosh [5]

Pasture-based and grass-fed branding are often associated with consumer perceptions of improved human health, environmental performance and animal welfare. Here, to examine the impacts of dairy production in detail, we contrasted global observational ($n = 156$) data for nitrogen and phosphorus losses from land by the duration of outdoor livestock grazing in confined, grazed and hybrid systems. Observational nitrogen losses for confined systems were lowest on a productivity—but not area—basis. No differences were noted for phosphorus losses between the systems. Modelling of the three dairy systems in New Zealand, the United States and the Netherlands yielded similar results. We found insufficient evidence that grazed dairy systems have lower nutrient losses than confined ones, but trade-offs exist between systems at farm scale. The use of a hybrid system may allow for uniform distribution of stored excreta, controlled dietary intake, high productivity and mitigation of animal welfare issues arising from climatic extremes.

An increasingly affluent global population is demanding more livestock-derived produce, but in tandem with improving environmental and animal welfare standards[1,2]. Demand for dairy products, in particular, has been globally forecast to increase by 1.6% per year between 2020 and 2029 (ref. [3]). Boosting international dairy production to meet this growing global demand will be achieved by a combination of the expansion of land used for dairying and increased intensification of existing dairying operations, alongside production efficiency gains[4]. However, there are growing concerns that some intensive dairy production systems may impair the environment by decreasing water quality[5,6], compromising animal welfare[7] and increasing greenhouse gas (GHG) emissions[8].

Dairying has been identified as a contributor to water quality deterioration in many jurisdictions[5,9–11]. Nutrient losses as nitrogen (N) and phosphorus (P) are important factors contributing to poor water quality via algal blooms and direct toxicity effects from nitrate N [12,13]. Literature reviews have shown that the annual loss of N and P from dairy farmland varies greatly from about 5 to 200 kg N ha^{-1} and from 0.5 to 10 kg P ha^{-1} (refs. [14–16]). This variation in contaminant loss is driven by contrasting climates, soils and landscapes, as well as on-farm management practice[17].

Over time, three general types of dairy production system have developed[14]. Grazing systems endeavour to match cow numbers with forage production on-farm and graze livestock outdoors most of the year (≥9 months). Hybrid systems use a combination of housing and enough land to support a few months (3–8) of grazing with forage preservation for non-growing periods. Confined systems house animals with most or all feed harvested and brought to the animals with minimal use of grazing (≤2 months).

Some studies, and consumer marketing campaigns, have suggested that grazing systems perform better in some regions than confined systems by certain 'sustainability' metrics, including animal health and welfare, reduced labour demands, profitability and nutrient losses to air and water, compared with confined systems[18–20]. In addition

---

[1]Lincoln Science Centre, AgResearch, Lincoln, New Zealand. [2]Faculty of Agriculture and Life Sciences, Lincoln University, Lincoln, New Zealand. [3]Agricultural Research Service, United States Department of Agriculture, University Park, PA, USA. [4]Agrosystems Research, Wageningen Plant Research, Wageningen, the Netherlands. [5]DairyNZ Ltd, Lincoln, New Zealand. ✉e-mail: richard.mcdowell@agresearch.co.nz

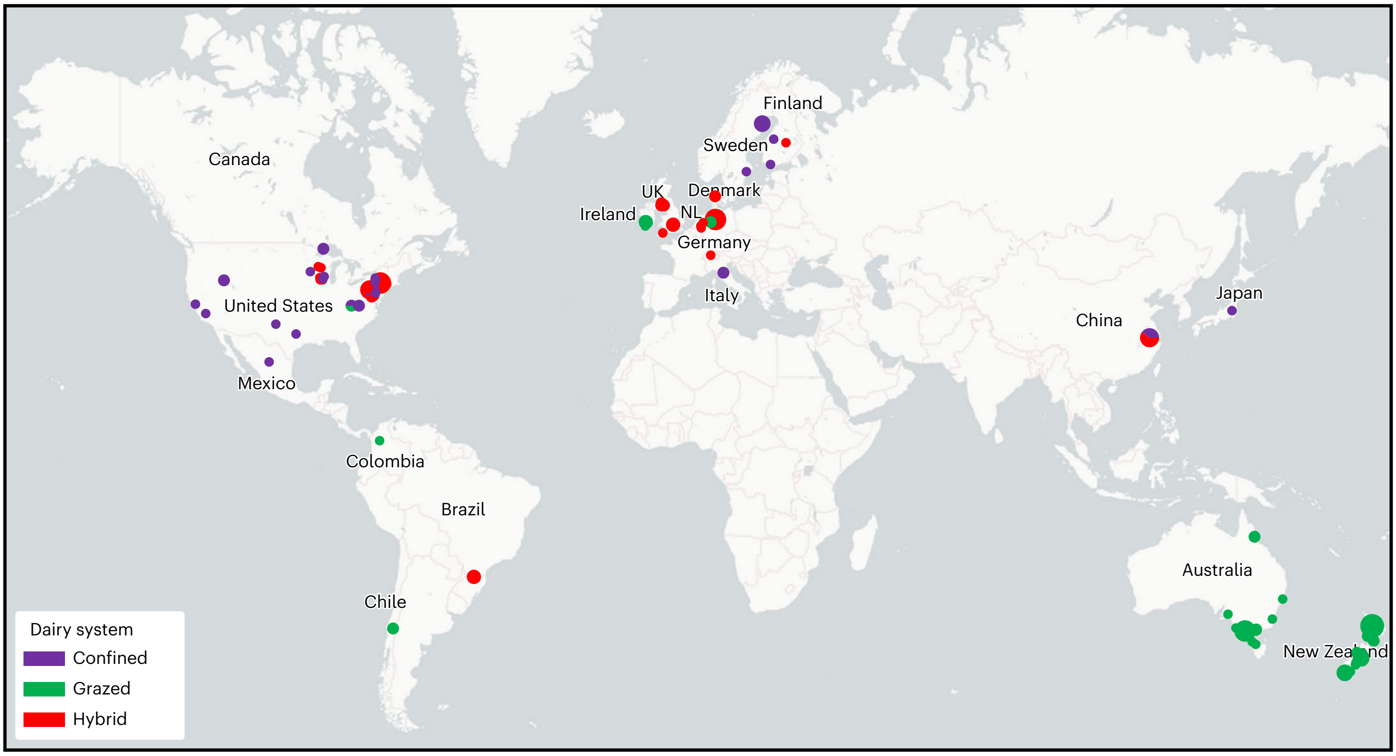

Fig. 1 | Location of the different dairy systems covered in the study. Data points used in the meta-analysis are indicated by coloured dots, including confined ($n = 32$), hybrid ($n = 49$) and grazed ($n = 75$) systems. Note that where data points are too close to be differentiated (<100 km), points are amalgamated, increasing the size of the mapped dots. The base map used data sourced from OpenStreetMap contributors available under an Open Database License (https://www.openstreetmap.org/copyright).

to some human health benefits, these metrics have been combined to infer that some grazed, grass-fed dairy systems are more 'environmentally and animal welfare friendly' than other systems from a consumer perspective[21], and therefore used to attract product premiums (https://www.origingreen.ie/) [22,23]. However, few empirical data exist to support these claims with respect to nutrient losses to water, especially in intensive systems where the potential for nutrient loss is greater than in systems with lower stocking rates[24]. Moreover, in a preliminary analysis conducted over a decade ago, Kleinman and Soder[14] hypothesized that hybrid systems may offer the ability to reduce nutrient losses by using housing to capture animal excreta and apply it to land to avoid storm events or wet times of the year when nutrient loss risks are high. Furthermore, hybrid and confined systems offer the ability to control and regulate dietary intake for higher milk production and the administration of feed additives to improve production or reduce enteric methane emission[25].

With an increased spotlight on dairying and nutrient losses, much more data are now available to compare systems. In this Analysis, hence, we aimed to determine whether differences exist between the three contrasting intensive dairy production systems based on the duration of outdoor grazing, with a focus on their N and P losses as indicators of water quality. However, we acknowledge that empirical data, from global studies, can be variable, being influenced by biophysical conditions and local management decisions. Therefore, we also modelled the likelihood of N and P losses from common dairy farming systems representative of those used in three major dairy-producing jurisdictions within the same terrestrial biome (New Zealand, the Netherlands, and the north-eastern United States). These were termed 'real' farms. We contrasted these real farms with hypothetical farms representative of the two less common systems to create outputs for the three production systems in each jurisdiction.

Results

Empirical observational data

Data were available to contrast N and P losses through runoff and leaching across the three different dairy systems for 156 sites (Supplementary Table 1) spread across 17 countries (Fig. 1) and 7 terrestrial biomes (Supplementary Table 2). Most data were available for New Zealand ($n = 45$), the United States ($n = 40$) and Australia ($n = 19$), with European countries (including the United Kingdom) accounting for 38 of the remaining sites (Supplementary Table 1). Across systems, most data were available for grazed ($n = 75$), followed by the hybrid ($n = 49$) and confined systems ($n = 32$). Across biomes, 104 sites were in the Temperate broadleaf forest biome, followed by 30 sites within the temperate grassland, savanna and shrubland biome (Supplementary Table 2). Sites were more evenly distributed within biomes for the confined systems than either the grazed or hybrid systems. For example, the temperate biomes accounted for 56% of sites for the confined system but 95% and 92% of the grazed and hybrid systems, respectively. Similarly, systems were more evenly spread in the United States and in Europe, whereas countries like Australia and New Zealand contained data exclusively from the grazed system.

Sufficient data were available to contrast the characteristics of each system (Table 1), although caution should be applied if contrasting the characteristics of systems with few data (for example, $n < 5$ for N surplus–confined, P fertilizer input–confined, and P surplus–grazed). Typically, confined systems were larger in size, stocked at a higher rate and had greater total N input to the soil (caused by manure not fertilizer) and N surplus, than hybrid or grazed systems. Nitrogen losses for confined systems were lower on a productivity basis (g kg^{-1} fat- and protein-corrected milk; FPCM) but not on an area basis. For P, P fertilizer and total P input were lower in confined than hybrid or grazed systems, but P losses on an area and productivity basis did not differ. Nitrogen

Table 1 | Descriptive statistics of the variables included in the study

Variable	System	N	N*	Median	Mean	s.e.m.	CV	Min	Max
Farm size (ha)	Grazed	75	0	30	401^{a}	100	217	<1	4,100
	Hybrid	46	3	64	248^{a}	153	418	1	6,900
	Confined	25	7	223	713^{b}	222	155	39	3,739
Stocking rate (cows ha^{-1} yr^{-1})	Grazed	59	16	3	2.9^{a}	0.1	33	1.0	5
	Hybrid	40	9	1.9	1.9^{b}	0.1	35	0.8	3.8
	Confined	25	7	2	3.3^{ab}	0.8	123	0.6	17.5
N input fertilizer (kg ha^{-1} yr^{-1})	Grazed	57	18	170	190^{b}	17	69	0	684
	Hybrid	27	22	150	145^{ab}	14	50	4	250
	Confined	18	14	93	116^{a}	20	73	16	280
Total N input (kg ha^{-1} yr^{-1})	Grazed	58	17	170	200^{a}	18	69	-	684
	Hybrid	31	18	170	174^{a}	18	58	20	465
	Confined	19	13	169	272^{a}	79	127	16	1,180
N surplus (kg ha^{-1} yr^{-1})	Grazed	16	59	311	261^{a}	28	42	−20	372
	Hybrid	22	27	132	155^{b}	20	59	7	385
	Confined	3	30	455	412^{ab}	248	85	164	660
N loss (kg ha^{-1} yr^{-1})	Grazed	63	12	35	44^{a}	4	78	4	175
	Hybrid	32	17	31	41^{a}	6	101	9	237
	Confined	24	8	42	106^{a}	29	136	7	486
N loss (g kg^{-1} FPCM yr^{-1})	Grazed	29	46	2.91	3.05^{a}	0.29	51	0.67	6.87
	Hybrid	18	31	2.70	4.17^{a}	0.77	78	1.24	12.37
	Confined	4	28	0.94	0.92^{b}	0.24	52	0.39	1.4
P input fertilizer (kg ha^{-1} yr^{-1})	Grazed	30	45	37	39^{b}	3	44	5	68
	Hybrid	14	35	14	27^{b}	7	101	5	90
	Confined	4	28	12	12^{a}	5	77	1	24
Total P input (kg ha^{-1} yr^{-1})	Grazed	29	46	40	42^{b}	3	36	15	68
	Hybrid	15	34	15	34^{b}	8	91	5	92
	Confined	6	26	18	18^{a}	3	34	10	25
P surplus (kg ha^{-1} yr^{-1})	Grazed	2	73	9.2	9.2^{a}	6.8	104	2.4	16
	Hybrid	16	33	3.4	10.1^{a}	4.4	173	−7.0	53.7
	Confined	8	24	1.5	2.8^{a}	1.3	129	−0.5	11
P loss (kg ha^{-1} yr^{-1})	Grazed	26	49	1.11	1.4^{a}	0.4	139	0.01	10
	Hybrid	26	23	1.21	4.3^{a}	2	242	0.02	50
	Confined	22	10	0.60	20.2^{a}	12.9	298	0.01	239
P loss (g kg^{-1} FPCM yr^{-1})	Grazed	7	68	0.09	0.09^{a}	0.02	71	0.01	0.17
	Hybrid	10	39	0.12	0.11^{a}	0.03	84	0.01	0.21
	Confined	5	27	0.14	0.37^{a}	0.22	134	0.03	1.2

For each, we include the number of data points present (*N*) and missing (*N**), mean, standard error of the mean (s.e.m.), coefficient of variation (CV), minimum (Min), median and maximum (Max) for stocking rate, and N and P inputs (fertilizer or fertilizer plus manure = total), surplus and loss in runoff and leaching by each system. All units are in kg (N or P) ha^{-1} yr^{-1} except stocking rate, which was recorded as across the farm. Means for dairy systems within a variable followed by the same letter are not different from one another (in log space) using Tukey's honestly significant difference at the $P < 0.05$ level.

loss per hectare was correlated to stocking rate and N surplus (Table 2). However, there were no substantial correlations between P inputs or P surplus and P losses per hectare. A paucity of significant correlations was expected given the diversity of farms within each system.

A paucity of data limited comparisons between all countries (Supplementary Table 1). However, sufficient data ($n > 5$) were available for Australia, Denmark, New Zealand and the United States. The analysis of these data indicated no differences in P losses, but a greater loss of N (g kg^{-1} FPCM) in Denmark (dominated by hybrid systems whose covered manure facilities and subsurface manure application reduce N losses as ammonia emissions to air over N losses to water[26]) compared with Australia (dominated by grazed systems) (Supplementary Table 3). Stocking rate was highest in New Zealand compared with the four other countries, as was median N surplus (which included N fixation by clover), but this was not reflected in N losses on neither a productivity nor an area basis. These data would suggest, on average, that there is no clear advantage of one system over another for reducing nutrient losses to water.

Modelled data

We used real and hypothetical farm characteristics to model and contrast N and P losses from the three different systems in representative

Table 2 | Spearman correlation coefficients between selected variables

	Stocking rate	P input fertilizer	Total P input	P surplus	P loss	P loss (FPCM)	N input fertilizer	Total N input	N surplus	N loss
P input fertilizer	0.459^{**}									
Total P input	0.550^{***}	0.894^{***}								
P surplus	0.036	0.685^{**}	0.774^{***}							
P loss	−0.144	−0.261	−0.163	−0.193						
P loss (FPCM)	−0.323	-0.664^{*}	-0.664^{*}	−0.336	0.932^{***}					
N input fertilizer	0.449^{***}	0.154	0.039	−0.336	−0.064	−0.090				
Total N input	0.451^{***}	0.142	0.119	−0.203	0.004	0.055	0.828^{***}			
N surplus	0.823^{***}	−0.293	−0.160	−0.332	0.108	−0.147	0.717^{***}	0.803^{***}		
N loss	0.284^{**}	0.159	0.019	0.225	−0.278	−0.342	0.082	0.171	0.369^{***}	
N loss (FPCM)	−0.122	−0.506	−0.506	0.550	−0.383	0.174	0.034	−0.097	−0.112	0.847^{***}

In this analysis we consider stocking rate (cow ha^{-1}) as well as N and P inputs (fertilizer or fertilizer plus manure indicates total, all in kg ha^{-1} yr^{-1}), surplus (kg ha^{-1} yr^{-1}) and loss by leaching and runoff (kg ha^{-1} yr^{-1} and g FPCM yr^{-1}). All units are in kg (N or P) ha^{-1} yr^{-1} except N and P losses as kg FPCM^{-1} and stocking rate, which was recorded as cow ha^{-1} yr^{-1} across the farm. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

dairying areas across three jurisdictions (varying in climate and soils). We used farm size, stocking rate and other characteristics (for example, soil types) representative of real farms found in each jurisdiction. The two hypothetical farms used the same farm size, stocking rate and characteristics but modified production according to typical practice for the system (Table 3).

As a check of model performance, we compared modelled and observed data (where available). Observed N losses were available for 3 years at the Manawatu site (8–21 kg N ha^{-1} yr^{-1}) (ref. [27]) and for 2 years at the Canterbury site (23–33 kg N ha^{-1} yr^{-1}) (ref. [28]). The modelled losses, averaged over 25 years of weather, compared well at 16 and 32 kg N ha^{-1} yr^{-1} for the Manawatu and Canterbury grazed farms, respectively. Modelled P losses averaged over 25 years for the Manawatu and Canterbury grazed farms were 1.68 and 0.42 kg ha^{-1} yr^{-1} (Table 4). No observations exist for P losses from these farms in the same years; however, measured losses in 2003 for the Manawatu farm were 1.68 kg P ha^{-1} (ref. [29]), and 0.25 kg P ha^{-1} yr^{-1} from 2001 to 2015 for the Canterbury farm[30], both of which are comparable to modelled outputs (Table 4). No data exist to compare across systems, and although the Integrated Farm System Model (IFSM) has been calibrated to a range of different systems in the United States and the Netherlands[8,31,32], we acknowledge that previous evaluation or calibration has not occurred for New Zealand systems.

In examining the different systems, N losses for most jurisdictions were lower from the hybrid or confined systems compared with the grazed system when expressed on a productivity basis (FPCM), ranging from a reduction of 0.5 g N kg^{-1} FPCM in Canterbury to 1.4 g N kg^{-1} FPCM in Pennsylvania (Table 4). The exception to this was an increase of 0.4 g N kg^{-1} FPCM for the hybrid compared with grazed system in the Netherlands. There was no consistent pattern in N losses across the three system types when expressed on an area basis. In contrast to N, P losses from grazed systems were generally lower than hybrid (mean reductions of 0.395 kg P ha^{-1} and 0.027 g P kg^{-1} FPCM) or confined (mean reductions of 0.766 kg P ha^{-1} and 0.038 g P kg^{-1} FPCM) systems on an area and productivity basis (Table 4). The exceptions were lower losses of 0.11 kg P ha^{-1} and 0.022 g P kg^{-1} FPCM for the hybrid system in Manawatu and 0.005 g kg^{-1} FPCM for the confined system in Pennsylvania.

As the amount of purchased grain varies across the three production systems and locations, another consideration is the nutrient losses that occur in the production of grain imported to the farm. Although we had no information of the production systems nor biophysical characteristics of these farms, we attempted to account for their contribution to nutrient losses. We used average crop yields from Table 3 and median N and P losses per hectare from Table 1 and multiplied these data by the amount of purchased grain in each system (Supplementary Information). This increased nutrient losses for production systems that required more purchased grain, which were primarily the confined feeding systems in New Zealand and all systems in the Netherlands. However, in general, the trends across production systems were like those found without considering the losses for purchased grain (Supplementary Table 6).

Discussion

Enhanced leaching and runoff N losses under grazing potentially reflect the non-uniform deposition of N-rich urine patches (600–1,000 kg N ha^{-1}), which exceed the nutrient requirements of pasture (200 kg N ha^{-1}) and are lost beyond the root zone, especially outside of the grass growing season[33]. Nitrogen losses from pasture systems, such as those in Australia and New Zealand, can be boosted by a high N surplus fed through a greater proportion of legume (20–40% clover) and therefore protein N in forage than pastures in, for example, the United States or Europe[8]. Uniquely, many grazed dairy farms in southern New Zealand also maintain a small area of crop (usually *Brassica*) used as forage in winter when pasture production is low. However, grazing these small areas in winter leads to urine being deposited onto the ground when nothing is growing, resulting in the mobilization of soil inorganic N and substantial N leaching (>100 kg N ha^{-1}) (ref. [34]).

In contrast to grazed systems, most N losses in hybrid and confined systems come from areas of the farm that are tilled and sown in crops (including off-farm losses associated with purchased grain), especially maize[8]. Other areas sown in alfalfa are not grazed in situ, but instead cut and fed in the housing system as part of a total mixed ration. Infrastructure within the housing system is used to capture and store excreta as solid or liquid manure that is then applied back to land as an organic nutrient source when soil moisture conditions are suitable, leaching risk is lower and plant growth is conducive to nutrient uptake[35]. This should in theory lower the potential for N losses to water on an area basis. However, no consistent pattern was noted on an area basis across the observations or the modelled farms. Instead, greater leaching losses from more highly fertilized tilled areas, that support hybrid and confined systems, potentially offset the lower urine patch deposition onto grazed pastures. However, total mixed rations, which are a common feeding strategy in housed systems, increase milk production per cow[23] and therefore reduce N losses on a productivity basis.

Contrasts between systems for P losses are less obvious than for N losses. This is caused by the strong influence of P loss processes such

Table 3 | Characteristics of farms modelled and evaluated using the ISFM

Parameter	Manawatu, New Zealand			Canterbury, New Zealand			Pennsylvania, United States			New York, United States			The Netherlands		
	Grazed (real)	Hybrid	Confined	Grazed (real)	Hybrid	Confined	Grazed	Hybrid	Confined (real)	Grazed	Hybrid	Confined (real)	Grazed	Hybrid (real)	Confined
Cow numbers	650	650	650	560	560	560	100	100	100	500	500	500	100	100	100
Farm area (ha)	255	255	255	285	285	285	100	100	100	530	530	530	63	63	63
Grass area (ha)	232	215	192	260	200	115	100	50	30	530	200	50	63	55	54
Grass yield (t DM ha^{-1})	13.6	13.4	13.4	17.4	17.5	17.4	7.3	6.6	6.3	8.7	8.3	6.0	9.6	9.7	10.4
Small grains or *Brassica* (ha)	23	40	63	25	85	170	–	–	–	–	–	–	–	–	–
Corn (ha)	–	–	–	–	–	–	–	30	50	–	180	280	–	8	9
Alfalfa (ha)	–	–	–	–	–	–	–	20	20	–	150	200	–	–	–
Farm N fertilizer (kg ha^{-1})	42	80	100	137	142	103	80	62	62	90	30	11	81	121	131
Farm P fertilizer (kg ha^{-1})	10	10	–	30	22	–	–	5	–	–	–	–	–	–	–
Replacement rate (%)	22	28	30	21	25	30	25	32	36	25	32	38	32	32	32
Milk production (kg FPCM cow^{-1})	6,021	7,000	9,000	6,377	7,500	8,600	6,000	8,700	10,000	7,600	9,500	11,500	7,916	9,755	11,114
Milk fat (%)	4.7	4.4	4.3	4.8	4.4	4.0	4.5	3.8	3.6	4.9	3.9	3.8	4.5	4.5	4.4
Grazed forage (t DM)	2,417	1,769	–	2,689	1,725	–	295	145	–	1,392	535	–	239	105	–
Harvested forage (t DM)	318	1,175	2,175	370	570	1,735	160	302	448	1,138	2,818	3,725	273	445	478
Harvested grain (t DM)	–	–	471	–	655	1,308	–	259	255	–	439	1,006	–	–	69
Imported grain (t DM)	–	658	1,111	51	241	239	57	–	–	1,043	214	–	120	237	196
Imported other feed (t DM)	920	544	911	337	450	664	50	7	69	150	674	666	131	48	124
Manure handling	Slurry	Slurry	Slurry	Slurry	Slurry	Slurry	Solid	Liquid	Liquid	Solid	Slurry	Liquid	Slurry	Slurry	Slurry
Soil P content	High	High	High	High	High	High	Very high	Very high	Very high	High	High	High	High	High	High

Additional weather and soil characteristics are available in Supplementary Tables 4 and 5. For farms in the Manawatu, Canterbury, Pennsylvania, New York, and the Netherlands, mean annual temperature was 13.4, 12.0, 12.1, 8.1 and 12.1 °C, respectively. Commensurate values of mean annual precipitation were 1,034, 613, 1,074, 1,001 and 1,074 mm and soil pH 5.8, 5.7, 6.5, 5.5 and 6.0. Soil texture for all farms was silt loam except for the clay loam in the farm from the Netherlands. Small grains are usually barley except for the grazed and hybrid systems in New Zealand where crops (usually *Brassica*) are grazed in situ to supplement winter feed. Grazed forage (t dry matter (DM) offered) consists of pasture and the grazed forage crops. Soil P content in New Zealand and the Netherlands is measured as Olsen P with high encompassing a range of 35–45 mg kg^{-1}. Soil P content in the United States was measured as Mehlich-3 P, with high and very high representing concentrations of 100–150 and >150 mg kg^{-1}, respectively.

as erosion and runoff that vary considerably in space and time. Our modelling suggested that P losses would be greater in the hybrid and confined systems than for a grazed system on an area basis (Table 4). This was driven by erosion of P from crop areas, but also by the rate and form of surface-applied manures, slurries and fertilizers[36]. However, the IFSM is not spatial[37]. While this does not matter for nitrate N, which is highly mobile and primarily lost by leaching, P can be retained either by sorption or filtration during runoff. Hence, P loss will be highly sensitive to the distance runoff must travel between being lost and entering the stream, as well as soil P retention capacities, and rates of soil erosion. Variation in P losses in grazed systems also comes from freshly grazed or senescent pasture, and from the treading action of grazing cattle that disturb the soil[38]. The presence of grazing animals in wet periods can greatly increase P and sediment losses[39,40], meaning that their location on a property, and distance away from a waterway, is another factor causing variation in the measurement of P losses. These factors help explain why contrasts in P losses were not detected in our observed data.

Implications for policy

Data suggesting a link between N and P surplus, and N and P loss, has been used to compare likely losses in Organisation for Economic Co-operation and Development and European Union countries and to guide water quality policy at the national and regional level, with some extrapolating this to the farm scale[41–43]. Policy aims to decrease N and P losses by reducing surpluses and P build-up in the soil[44,45]. For example, implementation of the European Union Water Framework Directive has halved the surplus of N since a peak of 250 kg N ha^{-1} in the mid-1990s by restricting the quantity of N that can be applied through manure and fertilizer[46]. However, as N and P are lost by different mechanisms that are expressed to different magnitudes in each dairy system, inter-country comparisons using nutrient surpluses might not give a true picture of nutrient losses to water, especially if countries are dominated by one system.

Supporting the use of an N surplus in policy, we found significant but weak correlations between observed N surplus, stocking rate and N loss across systems and locations. However, assuming the

Table 4 | Nitrogen and phosphorus combined runoff and leaching losses from each location and system per hectare on-farm area and per kilogram of FPCM, along with annual surplus N and P and GHG emissions

Location	System	N loss (kg ha^{1})	N loss (g kg^{-1} FPCM)	N surplus (kg ha^{-1})	P loss (kg ha^{1})	P loss (g kg^{-1} FPCM)	P surplus (kg P ha^{-1})	Annual GHG losses (t CO_2-eq ha^{-1})	Annual GHG losses (kg CO_2-eq kg $FPCM^{-1}$)
Manawatu, New Zealand	Grazed	16	1.1	146	1.68	0.110	15	16	1.04
	Hybrid	8	0.5	204	1.57	0.088	13	21	1.15
	Confined	12	0.5	268	3.45	0.150	28	26	1.12
Canterbury, New Zealand	Grazed	32	2.6	157	0.42	0.034	40	14	1.09
	Hybrid	27	1.8	200	0.56	0.038	26	16	1.06
	Confined	11	0.6	202	0.73	0.043	13	18	1.04
Pennsylvania, United States	Grazed	36	6.0	130	1.05	0.175	4	7	1.15
	Hybrid	41	4.7	110	1.76	0.202	1	8	0.92
	Confined	46	4.6	136	1.70	0.170	1	10	0.99
New York, United States	Grazed	35	4.8	122	0.15	0.021	8	6	0.86
	Hybrid	35	3.9	125	0.86	0.096	4	10	1.08
	Confined	39	3.7	123	1.23	0.116	2	9	0.89
Friesland, the Netherlands	Grazed	31	2.4	97	0.15	0.007	2	12	0.98
	Hybrid	43	2.8	121	0.15	0.009	2	14	0.91
	Confined	29	1.6	146	0.13	0.012	0	15	0.84

same total N application rate, soil type and climate, N losses resulting from the even spread of stored manure at times conducive to plant uptake should intuitively result in less N loss compared with losses from sporadic and concentrated urine patches in grazed systems. Our modelling of the different dairy systems also suggests the N surplus policy requires more interpretation. For instance, comparing the two most different systems (grazed and confined), our intuition was supported by the greater modelled N losses in the grazed over confined systems in New Zealand and the Netherlands. This did not happen in the US systems owing to the lower stocking rate (≤1 cow ha^{-1}, compared with 1.6 cows ha^{-1} in the Netherlands and 2.0–2.6 cows ha^{-1} in New Zealand) which probably decreased the influence of urine patches in N losses. However, an emphasis on N surplus is warranted if also considering emissions to air where high N losses from land to water in the grazed system in New Zealand and the Netherlands was balanced by lower ammonia emissions to air (Supplementary Fig. 1).

In contrast to N, P losses come from a wider array of sources. A focus on P surplus is understandable for a confined system where a P surplus will probably increase soil P concentrations and, if evenly distributed across the farm, P losses. Increased soil P is of particular concern as it can take many years to decline, thus sustaining high P losses for many decades in the interim[47]. However, while a P surplus could increase soil P concentrations in a grazed system, greater treading damage and stock excretal returns can account for the majority of farm P losses[48]. This means that, under the same P surplus, P losses from a grazed system have the potential to be lower, the same or greater than a confined system depending on if, for example, treading damage and excretal returns occur close to a stream. This variation in losses will be important to consider in policy decisions where the excessive growth of periphyton or phytoplankton in rivers and lakes is severely limited or co-limited by P or N (refs. [49,50]). When considering off-farm losses (Supplementary Table 6), there may be an opportunity to utilize spatial variation to purchase grain from areas known to have low potential for nutrient losses[43].

Wider implications

Most regions have derived a system that is profitable and therefore optimized, from the farmer's perspective to a particular geographical location, but some regions have the ability and the freedom to adopt multiple systems. There are many other drivers, in addition to policy, that influence the adoption of different dairy systems. When considering the modification of an existing system, it is important to consider wider implications such as animal welfare, efficiency and other environmental effects such as ammonia and GHG emissions, but especially improved profitability that may come from product premiums and lower on-farm costs for milk produced under grazing. In the United States, grass-fed organic milk attracts a substantial premium (approximately 100% in 2016 (ref. [51])) over milk produced from confined operations[52], similarly all grass-fed milk receives a premium of €0.50 per 100 kg of milk in the Netherlands[53]. New Zealand milk, and especially organic milk, gets a small premium (approximately 5–10% depending on year[54,55]) as dairy systems are almost entirely grazed and perceived by consumers to already capture many environmental and animal welfare benefits over confined systems overseas[1]. While these premiums can be linked to improved outcomes associated with many of these wider implications[18–20], from an N (and to a lesser extent P) loss perspective, we see insufficient evidence to substantiate this consumer perception. Moreover, on a productivity basis, we see evidence that N leaching and runoff losses from the farm are in fact greater from systems that incorporate grazing than not, because milk production is generally lower.

Global demand for milk products is growing. In areas concerned with nutrient losses, system choices may need to be based on environmental and economic metrics that are expressed on a productivity or area basis, and that include any off-farm losses from the production of purchased feed. Furthermore, grazing may be economically beneficial where the risk of nutrient loss is low and receiving waters are less susceptible to eutrophication, or where there is sufficient consumer concern over using confined systems that may result in poorer milk quality, animal welfare or enhanced GHG emissions (Table 4)[56].

Off-farm losses are spatially dependent and hence could be avoided if grain is purchased only from areas of low nutrient loss. To aid the consumer in choosing products with low nutrient losses, the productivity metric for on- and off-farm losses could be displayed. Some data have recently been produced for >57,000 food products sold in the United Kingdom, which incorporates eutrophication potential measured as P loss on an area basis. However, this does not consider losses from different dairy systems[57]. If producers chose a grazing system, there is some evidence to show that intensive grazing systems (for example, >3.5 cows ha^{-1}) can reduce their stocking rate (for example, to 2.2 cows ha^{-1}) and be more profitable[58]. However, even lowering stocking rates to 2.2 cows ha^{-1} may not be low enough to reduce N losses per hectare and per kilogram product below that of systems using housing. Perhaps the optimal system is a hybrid system that adopts partial housing during high-risk nutrient loss periods and has capacity to sustainably increase production using total mixed rations without increasing stocking rates[23]. If used, for example, in winter and early spring, partial housing could provide improved animal welfare outcomes[59] and capture and store excreta before being uniformly applied to land when plants are growing, and leaching and runoff are unlikely thereby reducing the likelihood of N and P losses[34]. Such trade-offs need to be considered for workable policy, and meeting consumer demand for dairy products with valid environmental and animal welfare claims, while maintaining viable farm businesses.

Methods

Data search and filters

We interrogated the SCOPUS and Web of Science databases and Google Scholar using combinations of N or P and runoff or leaching and dairy as search terms. We restricted our search to only those studies published in peer-reviewed journals between 1995 and 2020. We undertook to focus on empirical data or data generated from mechanistic-based and validated models of real farms. Focusing on the most recent 25 year period ensures that we obtained data characteristic of a generation of contemporary farming and farm practice. The use of empirical and validated models increases the likelihood that the data were generated with accuracy and precision[60]. In contrast, we excluded mass balance approaches. These approaches tend to calculate leaching and runoff losses by difference from inputs and all other outputs. Errors in inputs and outputs are therefore accumulated, meaning that losses can be highly uncertain. We also excluded studies of large catchments where dairy farm losses are likely to be diluted by other land uses or processes such as attenuation or lag times between on-farm practice and water quality response associated with convoluted flow pathways. Our filtered data therefore included:

- Current, unaltered and established (>3 years old) farm systems, excluding those that were optimized (for example, using a strategy to mitigate N or P losses);
- Studies whose data were obtained from scales ranging from large plot (0.1 ha) to catchments (2,000 ha), where catchments were dominated by ≥66% dairy land use. This recognized that, while it is a common perception that N loss is dominated by subsurface flow of nitrate and P loss is dominated by surface runoff of particulate P (ref. [12]), substantial losses of N and P occur via a mix of flow paths at different scales[61]; and
- Only mean annual losses as kg N or P $\mathrm{ha}^{-1}\,\mathrm{yr}^{-1}$, excluding losses that were identified by the authors as generated in years characterized by hydrological drought or floods.

We collected the following parameters: farm size, stocking rate, P fertility (as Olsen P concentration or via a conversion) and N and P inputs as feed, fertilizer and manure (including land-applied dirty water or dairy shed effluent), N or P surplus and N or P losses on an annual per hectare basis. Studies were further characterized by location (latitude and longitude), terrestrial biome[62] and dairy farm system class as grazed (grazing for ≥9 months), hybrid (grazed 3 and 8 months) or confined (grazed ≤2 months).

In addition to measured observations of N or P loss, we accepted modelled estimates of N and P loss (kg $\mathrm{ha}^{-1}\,\mathrm{yr}^{-1}$) from the following models that had been calibrated and verified for dairy farm systems: Annual P Loss Estimator (APLE)[63], Agricultural Policy Extender (APEX)[64], Agricultural Production Systems Simulator (APSIM)[65], Dairy Forage System Model (DAFOSYM)[66], DeNitrification-DeComposition (DNDC)[67,68], Groundwater Loading Effects of Agricultural Management Systems (GLEAMS)[69], Integrated Farm System Model (IFSM)[37], Nitrate Leaching and Economic Analysis Package (NLEAP)[70] and OVERSEER[71]. We also calculated N and P losses, standardized to 4% fat and 3.4% protein as grams of N or P lost per kilogram of FPCM as per the International Dairy Federation[72].

Data analyses

General descriptive statistics were generated for each parameter. Owing to the highly skewed nature of most data, values were log-transformed before contrasting parameter means across system classes by an analysis of variance and the results presented as Tukey's honestly significant difference. To show the strength of associations between parameters, Spearman correlation coefficients were also generated.

Modelling

Additional information was obtained for the likelihood of N and P losses across the three different systems in New Zealand (Manawatu and Canterbury), the north-eastern United States (Pennsylvania and New York) and the Netherlands (Friesland) through modelling. As dairy production systems have generally been developed and attuned to a geographical region, one of the systems often dominates in each region but the others may be found, particularly for specialty markets such as organic and grass-fed milk. With some exclusions (for example, grazing in very cold climates), farm management and infrastructure have developed sufficiently to enable each system to be used, across many jurisdictions. To test the environmental performance of each system, in relation to N and P losses, and within an area where they are realistically possible, we chose to model examples of current systems within one terrestrial biome. The terrestrial biome chosen was 'temperate broadleaf forest', which represents climate and soil characteristics of dairy-intensive regions of New Zealand, the Netherlands and the north-eastern United States.

ISFM

Dairy production systems were simulated in each region using the IFSM software tool[37] (additional details about the model are given in Supplementary Information). Of the modelling candidates listed above, IFSM has been well used in modelling nutrient transformation and loss across all our systems[73–75]. This whole-farm model simulates crop production, feed use and the return of manure nutrients back to the land for up to 25 years of daily weather. Daily growth and development of crops are predicted on the basis of soil water and N availability, ambient temperature and solar radiation. Simulated tillage, planting, harvest, storage and feeding operations predict resource use, timeliness of operations, crop losses and nutritive quality of feeds produced.

Nutrient flows are tracked to predict losses to the environment and potential accumulation in the soil. Losses include ammonia volatilization, nitrification, denitrification and leaching losses of N, erosion of sediment, and runoff of sediment-bound and dissolved N and P across the farm boundaries. Carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O) emissions are tracked from soil, manure and machinery sources and sinks to predict cradle to farm-gate GHG emission in CO_2-eq units. Whole-system balances of N, P and carbon (C) are determined as the sum of nutrient imports in purchased feed, manure, fertilizer, deposition and fixation, minus the nutrient exports in milk, animals and feeds sold or lost.

Dairy production systems

Representative dairy production systems were simulated at two locations in New Zealand (Canterbury and Manawatu), two in the north-eastern United States. (Pennsylvania and New York) and one in the Netherlands (Table 3). At each location, the three production systems were modelled and evaluated: grazing, full confinement and hybrid (semi-confinement). With grazing systems, cattle had access to pastureland for at least 9 months of the year with low use of housing facilities and associated manure handling. Minimal feed supplementation was used to maintain a relatively low milk production (Table 3). Housing facilities normally consisted of an open lot or bedded pack barn, but in the Netherlands, a slatted floor barn was used. For full-confinement systems, cows were maintained all year in free-stall barns and fed total mixed rations to meet their nutrient needs for a relatively high milk production. For the hybrid systems, animals were housed in barns for about half of the year with use of grazing during the warmer portion of the year. Feed supplementation and milk production fell between that of the grazing and full-confinement systems.

Model parameters for the various production systems were set on the basis of the published studies in each region[8,27,28,76]. Data used from these studies included animal production (for example, stock numbers, cow body weight, milk production, feeding method, dietary protein and P), soil conditions (for example, soil type, soil P fertility, texture and pH), crop and pastureland (area, species mix, fertilizer and manure application rates, months grazed and yield), levels of supplementary feed and manure handling systems. Phosphorus fertilizer use in the hybrid and confined systems were often set to zero in the United States and the Netherlands examples (Table 3). This fertilizer strategy was afforded by high imported P in feed and adequate to high soil P concentrations where local regulations and guidelines may limit P application[44,45]. We did not set P fertilizer use to zero in the grazed or hybrid systems in New Zealand owing to evidence that at modest soil P concentrations stopping P fertilizer applications can decrease pasture production by 10–20% in 4–5 years (refs. [77,78]).

Each production system was simulated over 25 years of historical weather for the location. The same soil characteristics were used for each production system at a location where the soil texture, bulk density and available water holding capacity were set to represent the soil at the location. Feed crops produced were selected from those produced in the region that best met the nutrient needs of the animals. Crops varied from all grass (plus clover) for grazing systems to a mixture of grass, legume, grain-crop silage and grain crops with confinement systems (Table 3). Farm management parameters such as crop and pasture yields, planting and harvest dates, equipment used, tillage practices, feed storage and manure handling methods were representative of those in each region.

Herd size and land area were held the same across the three production strategies at each region. Herd size was set representative of that found in the region (Table 3). Herds in New Zealand were only cows, while those in the United States and the Netherlands included replacement heifers. Crops and land area used for each crop varied to best meet the feed needs of the herd (Table 3). Feeds were purchased as needed to supplement that produced on the farm to meet the nutrient needs of all animal groups making up the herd, which brought additional nutrients onto the farm. When more feed was produced than needed by the herd, that feed was sold and the nutrients contained were exported from the farm. The authenticity of these representative farms was checked with local dairy farm consultants and colleagues.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

We used the SCOPUS database to source data for our observations. The observational and modelled data used or generated in this study are available at figshare (https://doi.org/10.6084/m9.figshare.20486373). Source data are provided with this paper.

Code availability

The statistical coding is available from the corresponding author on reasonable request.

References

1. Alonso, M. E., González-Montaña, J. R. & Lomillos, J. M. Consumers' concerns and perceptions of farm animal welfare. *Animals* **10**, 385 (2020).
2. Bai, Z. et al. China's livestock transition: driving forces, impacts, and consequences. *Sci. Adv.* **4**, eaar8534 (2018).
3. OECD & FAO *OECD-FAO Agricultural Outlook 2020–2029* (OECD, 2020); https://www.oecd-ilibrary.org/content/publication/1112c23b-en
4. Alexander, P. et al. Losses, inefficiencies and waste in the global food system. *Agric. Syst.* **153**, 190–200 (2017).
5. Scarsbrook, M. R. & Melland, A. R. Dairying and water-quality issues in Australia and New Zealand. *Anim. Prod. Sci.* **55**, 856–868 (2015).
6. Clark, M. & Tilman, D. Comparative analysis of environmental impacts of agricultural production systems, agricultural input efficiency, and food choice. *Environ. Res. Lett.* **12**, 064016 (2017).
7. Mee, J. F. & Boyle, L. A. Assessing whether dairy cow welfare is "better" in pasture-based than in confinement-based management systems. *N. Z. Vet. J.* **68**, 168–177 (2020).
8. Rotz, C. A., Holly, M., de Long, A., Egan, F. & Kleinman, P. J. A. An environmental assessment of grass-based dairy production in the northeastern United States. *Agric. Syst.* **184**, 102887 (2020).
9. Julian, J. P., de Beurs, K. M., Owsley, B., Davies-Colley, R. J. & Ausseil, A. G. E. River water quality changes in New Zealand over 26 years: response to land use intensity. *Hydrol. Earth Syst. Sci.* **21**, 1149–1171 (2017).
10. Glibert, P. M. From hogs to HABs: impacts of industrial farming in the US on nitrogen and phosphorus and greenhouse gas pollution. *Biogeochemistry* **150**, 139–180 (2020).
11. Hoekstra, N. J. et al. Scenarios to limit environmental nitrogen losses from dairy expansion. *Sci. Total Environ.* **707**, 134606 (2020).
12. Carpenter, S. R. et al. Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecol. Appl.* **8**, 559–568 (1998).
13. Camargo, J. A., Alonso, A. & Salamanca, A. Nitrate toxicity to aquatic animals: a review with new data for freshwater invertebrates. *Chemosphere* **58**, 1255–1267 (2005).
14. Kleinman, P. J. A. & Soder, K. in *Environmental Impacts of Pasture-Based Farming* (ed. McDowell, R. W.) Ch. 10 (CABI, 2008).
15. Wilcock, R. J. et al. Trends in water quality of five dairy farming streams in response to adoption of best practice and benefits of long-term monitoring at the catchment scale. *Mar. Freshw. Res.* **64**, 401–412 (2013).
16. Davies-Colley, R. J. & Nagels, J. W. Effects of dairying on water quality of lowland streams in Westland and Waikato. *Proc. N. Z. Grass. Assoc.* **64**, 107–114 (2002).
17. McDowell, R. W. & Wilcock, R. J. Water quality and the effects of different pastoral animals. *N. Z. Vet. J.* **56**, 289–296 (2008).
18. Gilker, R. E. & Weil, R. R. Inorganic nitrogen losses to groundwater are minimal from two management-intensive grazing dairy farms in Maryland. *Renew. Agric. Food Syst.* **33**, 347–359 (2018).
19. O'Brien, D. et al. A life cycle assessment of seasonal grass-based and confinement dairy farms. *Agric. Syst.* **107**, 33–46 (2012).
20. Pasture-Fed Livestock Association et al. *The Animal Welfare and Environmental Benefits of Pasture for Life Farming—Interim Findings* (Pasture for Life, 2018); https://www.agricology.co.uk/sites/default/files/the-animal-welfare-and-environmental-benefits-of-pasture-for-life-farming.pdf
21. Johnson, D. M. et al. Sustainability of management-intensive grazing dairy farms versus conventional confinement dairy farms. In *Proc. 11th European IFSA Symposium, Farming Systems Facing Global Challenges: Capacities and Strategies* (eds Aenis, T. et al.) 1424–1431 (IFSA, 2014).

22. Synlait Milk Ltd *Lead with Pride: Certified Best Practice Dairy Farming* (Synlait, 2018); https://www.synlait.com/wp-content/uploads/2018/10/Lead-With-Pride-Brochure.pdf
23. Joubran, A. M., Pierce, K. M., Garvey, N., Shalloo, L. & O'Callaghan, T. F. Invited review: A 2020 perspective on pasture-based dairy systems and products. *J. Dairy Sci.* **104**, 7364–7382 (2021).
24. Beukes, P. C., Romera, A. J., Neal, M. & Mashlan, K. Performance of pasture-based dairy systems subject to economic, climatic and regulatory uncertainty. *Agric. Syst.* **174**, 95–104 (2019).
25. Ferris, C. P., Watson, S., Gordon, A. W. & Barley, J. Physical and economic performance of dairy cows managed within contrasting grassland-based milk production systems over 3 successive lactations. *J. Dairy Sci.* **105**, 3153–3175 (2022).
26. Sorensen, C. A. G., Sommer, S. G., Bochtis, D. & Rotz, A. in *Animal Manure Recycling: Treatment and Management* (eds Sommer, S. G. et al.) Ch. 12 (Wiley, 2013).
27. Christensen, C. L., Hedley, M. J., Hanly, J. A. & Horne, D. J. Duration-controlled grazing of dairy cows. 2: nitrogen losses in sub-surface drainage water and surface runoff. *N. Z. J. Agric. Res.* **62**, 48–68 (2019).
28. Chapman, D. F. et al. Production, profit and nitrogen flows in irrigated dairy systems representing different industry development pathways: the Pastoral 21 experience in Canterbury. *N. Z. J. Agric. Res.* **64**, 3–35 (2021).
29. Houlbrooke, D. J., Horne, D. J., Hedley, M. J., Snow, V. & Hanly, J. A. Land application of farm dairy effluent to a mole and pipe drained soil: implications for nutrient enrichment of winter-spring drainage. *Aust. J. Soil Res.* **46**, 45–52 (2008).
30. McDowell, R. W., Gray, C. W., Cameron, K. C., Di, H. J. & Pellow, R. The efficacy of good practice to prevent long-term leaching losses of phosphorus from an irrigated dairy farm. *Agric. Ecosyst. Environ.* **273**, 86–94 (2019).
31. Rotz, C. A., Oenema, J. & van Keulen, H. Whole farm management to reduce nutrient losses from dairy farms: a simulation study. *Appl. Eng. Agric.* **22**, 773–784 (2006).
32. Veltman, K. et al. Comparison of process-based models to quantify nutrient flows and greenhouse gas emissions associated with milk production. *Agric. Ecosyst. Environ.* **237**, 31–44 (2017).
33. Di, H. J. & Cameron, K. C. Nitrate leaching losses and pasture yields as affected by different rates of animal urine nitrogen returns and application of a nitrification inhibitor—a lysimeter study. *Nutr. Cycling Agroecosyst.* **79**, 281–290 (2007).
34. Monaghan, R. M., Laurenson, S., Dalley, D. E. & Orchiston, T. S. Grazing strategies for reducing contaminant losses to water from forage crop fields grazed by cattle during winter. *N. Z. J. Agric. Res.* **60**, 333–348 (2017).
35. Liu, J. et al. A review of regulations and guidelines related to winter manure application. *Ambio* **47**, 657–670 (2018).
36. Spiegal, S. et al. Manuresheds: advancing nutrient recycling in US agriculture. *Agric. Syst.* **182**, 102813 (2020).
37. Rotz, A. *The Integrated Farm System Model: Reference Manual Version 4.4. 254* (USDA, 2018); https://www.ars.usda.gov/ARSUserFiles/80700500/Reference%20Manual.pdf
38. McDowell, R. W., Nash, D., George, A., Wang, Q. J. & Duncan, R. Approaches for quantifying and managing diffuse phosphorus exports at the farm/small catchment scale. *J. Environ. Qual.* **38**, 1968–1980 (2009).
39. Monaghan, R. M., Smith, L. C. & Muirhead, R. W. Pathways of contaminant transfers to water from an artificially-drained soil under intensive grazing by dairy cows. *Agric. Ecosyst. Environ.* **220**, 76–88 (2016).
40. Donovan, M. Modelling soil loss from surface erosion at high-resolution to better understand sources and drivers across land uses and catchments; a national-scale assessment of Aotearoa, New Zealand. *Environ. Model. Software* **147**, 105228 (2022).
41. *Measuring the Environmental Performance of Agriculture Across OECD Countries* (OECD, 2021).
42. Klages, S. et al. Nitrogen surplus—a unified indicator for water pollution in Europe? *Water* **12**, 1197 (2020).
43. Quemada, M. et al. Exploring nitrogen indicators of farm performance among farm types across several European case studies. *Agric. Syst.* **177**, 102689 (2020).
44. van Leeuwen, M. M. W. J. et al. The relevance of spatial scales in nutrient balances on dairy farms. *Agric. Ecosyst. Environ.* **269**, 125–139 (2019).
45. Chapman, D. et al. Production, profit and nitrogen flows in irrigated dairy systems representing different industry development pathways: the Pastoral 21 experience in Canterbury. In *Proc. Australasian Dairy Science Symposium* 187–190 (Australasian Dairy Science Symposium Limited, 2018).
46. Van Grinsven, H. J. M., Tiktak, A. & Rougoor, C. W. Evaluation of the Dutch implementation of the nitrates directive, the water framework directive and the national emission ceilings directive. *NJAS Wagening. J. Life Sci.* **78**, 69–84 (2016).
47. McDowell, R., Dodd, R., Pletnyakov, P. & Noble, A. The ability to reduce soil legacy phosphorus at a country scale. *Front. Environ. Sci.* https://doi.org/10.3389/fenvs.2020.00006 (2020).
48. McDowell, R. W., Nash, D. M. & Robertson, F. Sources of phosphorus lost from a grazed pasture receiving simulated rainfall. *J. Environ. Qual.* **36**, 1281–1288 (2007).
49. Arvola, L., Järvinen, M. & Tulonen, T. Long-term trends and regional differences of phytoplankton in large Finnish lakes. *Hydrobiologia* **660**, 125–134 (2011).
50. Watson, S. B. et al. The re-eutrophication of Lake Erie: harmful algal blooms and hypoxia. *Harmful Algae* **56**, 44–66 (2016).
51. Organic costs and returns: milk production costs and returns from the 2016 ARMS. *USDA* https://www.ers.usda.gov/data-products/commodity-costs-and-returns/organic-costs-and-returns/ (2021).
52. Wong, J., Raghunathan, U., Escalante, C. L. & Wolfe, K. Consumer premiums for environmentally friendly grass-fed and organic milk in the Southeast. *J. Agribus.* **28**, 75–88 (2010).
53. Milk price system. *FrieslandCampina* https://www.frieslandcampina.com/our-farmers/owned-by-farmers/milk-price-system/ (2022).
54. Kissun, S. Organic milk price dropped. *Rural News Group* https://www.ruralnewsgroup.co.nz/dairy-news/dairy-general-news/organic-milk-price-dropped (2021).
55. Fonterra expands its organic milk business to meet growing consumer appetite. *Fonterra Co-op Ltd* https://www.fonterra.com/nz/en/our-stories/media/fonterra-expands-its-organic-milk-business-to-meet-growing-consumer-appetite.html (2015).
56. Mazzetto, A., Falconer, S. & Ledgard, S. Mapping the carbon footprint of milk for dairy cows. *AgResearch* https://www.dairynz.co.nz/media/5794083/mapping-the-carbon-footprint-of-milk-for-dairy-cows-report-updated.pdf (2021).
57. Clark, M. et al. Estimating the environmental impacts of 57,000 food products. *Proc. Natl Acad. Sci. USA* **119**, e2120584119 (2022).
58. MacDonald, K. A., Beca, D., Penno, J. W., Lancaster, J. A. S. & Roche, J. R. Short communication: effect of stocking rate on the economics of pasture-based dairy farms. *J. Dairy Sci.* **94**, 2581–2586 (2011).
59. Blanco-Penedo, I., Ouweltjes, W., Ofner-Schröck, E., Brügemann, K. & Emanuelson, U. Symposium review: animal welfare in free-walk systems in Europe. *J. Dairy Sci.* **103**, 5773–5782 (2020).

60. Pellerin, B. A. et al. Emerging tools for continuous nutrient monitoring networks: sensors advancing science and water resources protection. *J. Am. Water Resour. Assoc.* **52**, 993–1008 (2016).
61. Woodward, S. J. R. & Stenger, R. Bayesian chemistry-assisted hydrograph separation (BACH) and nutrient load partitioning from monthly stream phosphorus and nitrogen concentrations. *Stoch. Environ. Res. Risk Assess.* **32**, 3475–3501 (2018).
62. Dinerstein, E. et al. An ecoregion-based approach to protecting half the terrestrial realm. *Bioscience* **67**, 534–545 (2017).
63. Vadas, P. A., Mark Powell, J., Brink, G. E., Busch, D. L. & Good, L. W. Whole-farm phosphorus loss from grazing-based dairy farms. *Agric. Syst.* **140**, 40–47 (2015).
64. Mason, R. E., Görres, J., Faulkner, J. W., Doro, L. & Merrill, S. C. Calibrating the apex model for simulations of environmental and agronomic outcomes on dairy farms in the Northeast U.S.: a step-by-step example. *Appl. Eng. Agric.* **36**, 281–301 (2020).
65. Holzworth, D. et al. APSIM Next Generation: overcoming challenges in modernising a farming systems model. *Environ. Model. Software* **103**, 43–51 (2018).
66. Rotz, C. A., Black, J. R., Mertens, D. R. & Buckmaster, D. R. DAFOSYM: a model of the dairy forage system. *J. Prod. Agric.* **2**, 83–91 (1989).
67. Li, C., Frolking, S. & Frolking, T. A. A model of nitrous oxide evolution from soil driven by rainfall events: 1. Model structure and sensitivity. *J. Geophys. Res. Atmosph.* **97**, 9759–9776 (1992).
68. Deng, M. et al. Evaluation of N environmental risks on Andosols from an intensive dairy farming watershed using DNDC. *Sci. Total Environ.* **512–513**, 659–671 (2015).
69. Leonard, R. A., Knisel, W. G. & Still, D. A. GLEAMS: groundwater loading effects of agricultural management systems. *Trans. ASAE* **30**, 1403–1418 (1987).
70. Shaffer, M. J., Halvorson, A. D. & Pierce, F. J. in *Managing Nitrogen for Groundwater Quality and Farm Profitability* (eds Follett, R. F. et al.) Ch. 13 (Soil Science Society of America, 1991).
71. OVERSEER Nutrient Budgets v.6.2.2. (AgResearch, 2016).
72. International Dairy Federation *A Common Carbon Footprint Approach for the Dairy Sector: The IDF Guide to Standard Life Cycle Assessment Methodology* (International Dairy Federation, 2015).
73. Ma, L. et al. Application of grazing land models in ecosystem management: current status and next frontiers. *Adv. Agron.* **158**, 173–215 (2019).
74. Deak, A., Hall, M. H., Sanderson, M. A., Rotz, A. & Corson, M. Whole-farm evaluation of forage mixtures and grazing strategies. *Agron. J.* **102**, 1201–1209 (2010).
75. Veltman, K. et al. Assessing and reducing the environmental impact of dairy production systems in the northern US in a changing climate. *Agric. Syst.* **192**, 103170 (2021).
76. Oenema, J. & Oenema, O. Intensification of grassland-based dairy production and its impacts on land, nitrogen and phosphorus use efficiencies. *Front. Agr. Sci. Eng.* **8**, 130–147 (2021).
77. McBride, S. D., Nguyen, M. L. & Rickard, D. S. Implications of ceasing annual superphosphate topdressing applications on pasture production. *Proc. N. Z. Grass. Assoc.* **52**, 177–180 (1990).
78. O'Connor, M. B., Smart, C. E. & Ledgard, S. F. Long term effects of withholding phosphate applicaton on North Island hill country: Te Kuiti. *Proc. N. Z. Grass. Assoc.* **51**, 21–24 (1990).

Acknowledgements

R.W.M. acknowledges the support of the Our Land and Water National Science Challenge (New Zealand Ministry of Business, Innovation and Employment grant number C10X1507; R.W.M.) while writing the paper. We appreciate the constructive reviews from colleagues. The views presented are those of the authors and not of their constituent institutions.

Author contributions

R.W.M conceived the idea, compiled and analysed the observational data and wrote the paper. C.A.R. undertook the modelling and helped write the paper. J.O. and K.A.M. helped source observational data, provided farm system data for modelling and helped write the paper. All authors contributed with discussion and critically reviewed the manuscript.

Competing interests

K.A.M. is employed by DairyNZ Ltd, an organization funded by levies from dairy farmers in New Zealand. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43016-022-00644-2.

Correspondence and requests for materials should be addressed to R. W. McDowell.

Peer review information *Nature Food* thanks Will Brownlie, Perttu Virkajärvi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Corresponding author(s):	R.W. MCDOWELL
Last updated by author(s):	Sep 27, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	*No software was used.*
Data analysis	The analysis involved log-transformation and an analysis of variance with a post hoc test (Tukey's HSD). No bespoke code was used to do this. For dairy system modelling we used the Integrated Farm System model version 4.6. The input and output files for the IFSM are available at the Figshare repository.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Please see the data availability statement. The data are also supplied in the Supplementary Information and at Figshare: https://doi.org/10.6084/m9.figshare.20486373

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	*NA*
Population characteristics	*NA*
Recruitment	*NA*
Ethics oversight	*NA*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	*NA*
Data exclusions	*NA*
Replication	*NA*
Randomization	*NA*
Blinding	*NA*

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*NA*
Research sample	*NA*
Sampling strategy	*NA*
Data collection	*NA*
Timing	*NA*
Data exclusions	*NA*
Non-participation	*NA*
Randomization	*NA*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*The study was an analysis of existing literature, classifying and contrasting dairy systems with different periods of grazing.*
Research sample	*Data were sourced from published studies and classed into grazed, partial-housing (hybrid) and confined, corresponding to nine or more, between two and eight and one or two months of grazing, respectively.*
Sampling strategy	*We sourced data from studies published in peer-reviewed journals between 1995 and 2020 which ensures that we obtained data characteristic of a generation of contemporary farming and farm practice. To increase the likelihood that the data were generated with accuracy and precision, we focused on empirical data or data generated from mechanistic-based and validated models of real farms.*
Data collection	*We interrogated the SCOPUS and Web of Science databases, and Google Scholar using combinations of N or P and runoff or leaching and dairy as search terms.*
Timing and spatial scale	*Please see the sampling strategy.*
Data exclusions	*We excluded grey literature, studies older than 1995, those studies using mass balance approaches and studies of large catchments where dairy farm losses are likely to be diluted by other land uses or processes like attenuation or lag times between on-farm practice and water quality response.*
Reproducibility	*Our filtered data included:* *• Current, unaltered and established (> 3 years old) farm systems, excluding those that were optimised (e.g., using a strategy to mitigate N or P losses);* *• Studies whose data were obtained from scales ranging from large plot (0.1ha) to catchments (2000 ha) - where catchments were dominated by ≥ 66% dairy land use; and* *• Only mean annual losses as kg N or P/ha/yr, excluding losses that were identified by the authors as generated in years characterised by hydrological drought or floods.* *In addition to measured observations of N or P loss ,we accepted modelled estimates of N and P loss (kg ha-1 yr-1) from the following models that had been calibrated and verified for dairy farm systems: Annual P Loss Estimator (APLE); Agricultural Policy Extender (APEX); Agricultural Production Systems Simulator (APSIM); Dairy Forage System Model (DAFOSYM);; DeNitrification-DeComposition (DNDC); Groundwater Loading Effects of Agricultural Management Systems (GLEAMS); Integrated Farm System Model (IFSM); Nitrate Leaching and Economic Analysis Package (NLEAP); and OVERSEER.*
Randomization	*Please see "Research sample". In addition to data for N and P loss we also collected data for farm size, stocking rate, P fertility, N and P inputs as feed, fertiliser, and manure (including land-applied dirty water or dairy shed effluent), and N or P surplus. We further characterised studies by location (latitude and longitude), terrestrial biome. These data were compared to N and P losses and also used to contrast N and P losses by jurisdiction.*
Blinding	*Blinding was not applicable to this study. We describe our filtering rules in detail to constrain the data and avoid bias as much as possible.*

Did the study involve field work? ☐ Yes ☒ No

Field work, collection and transport

Field conditions	*NA*
Location	*NA*
Access & import/export	*NA*
Disturbance	*NA*

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a | Involved in the study
☒ ☐ Antibodies
☒ ☐ Eukaryotic cell lines
☒ ☐ Palaeontology and archaeology
☒ ☐ Animals and other organisms
☒ ☐ Clinical data
☒ ☐ Dual use research of concern

Methods

n/a | Involved in the study
☒ ☐ ChIP-seq
☒ ☐ Flow cytometry
☒ ☐ MRI-based neuroimaging

Antibodies

Antibodies used	*Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Validation	*Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.*

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	*State the source of each cell line used and the sex of all primary cell lines and cells derived from human participants or vertebrate models.*
Authentication	*Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.*
Mycoplasma contamination	*Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.*
Commonly misidentified lines (See ICLAC register)	*Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

Palaeontology and Archaeology

Specimen provenance	*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.*
Specimen deposition	*Indicate where the specimens have been deposited to permit free access by other researchers.*
Dating methods	*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	*For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Reporting on sex	*Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.*

Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☒	☐	Public health
☒	☐	National security
☒	☐	Crops and/or livestock
☒	☐	Ecosystems
☒	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☒	☐	Demonstrate how to render a vaccine ineffective
☒	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☒	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☒	☐	Increase transmissibility of a pathogen
☒	☐	Alter the host range of a pathogen
☒	☐	Enable evasion of diagnostic/detection modalities
☒	☐	Enable the weaponization of a biological agent or toxin
☒	☐	Any other potentially harmful combination of experiments and agents

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	*For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*
Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session (e.g. UCSC)	

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates	*Describe the experimental replicates, specifying number, type and replicate agreement.*
Sequencing depth	*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*
Antibodies	*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Peak calling parameters	*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*
Data quality	*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*
Software	*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*
Instrument	*Identify the instrument used for data collection, specifying make and model number.*
Software	*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*
Cell population abundance	*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*
Gating strategy	*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*
Volume censoring	*Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings	*Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*
Effect(s) tested	*Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016)	*Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*
Correction	*Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

Models & analysis

n/a	Involved in the study	
☐	☐	Functional and/or effective connectivity
☐	☐	Graph analysis
☐	☐	Multivariate modeling or predictive analysis

Functional and/or effective connectivity	*Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*
Graph analysis	*Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*
Multivariate modeling and predictive analysis	*Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*