Supplementary Material

Bayesian belief network analysis infers the importance of post-construction support in maintaining the functionality of pit latrines and septic systems across 12 countries

List of tables

List of figures

Table S1. Variables used in the Bayesian Belief Network of sanitation system functionality and scientific justification for key O&M tasks.

Nodes (short name)[a]	Description of variable	Output states	Rationale and Evidence
overflowing	Indicated by evidence of a sanitation facility being full and allowing waste to be discharged onto the ground at the time of the survey.	no yes	N/A
leaking	Indicated by evidence of wastewater escaping from the sanitation facility through structural defects, such as cracks, gaps, or faulty seals, at the time of the survey, regardless of whether the system was full.	no yes	N/A
floodrisk	Represents the likelihood of flooding based on the historical frequency of floods in the area.	low-risk medium-risk high-risk	Floodwater can damage the superstructures of pit latrines and septic systems, leading to overflows. Additionally, floods may hinder essential maintenance activities by causing delays in emptying, restricting access to supplies and spare parts, and potentially increasing maintenance costs (Hyde-Smith et al., 2022).
facilityfee	Indicates whether users pay a fee to use a facility.	no yes	The operation and maintenance of many sanitation systems and fecal sludge treatment rely entirely on direct user fees, which can help cover infrequent costs such as spare parts, repairs, and desludging (World Health Organization, 2018a).
cleaningperson	Indicates whether a person – formally employed or informally engaged - is responsible for cleaning the facility.	no yes	A dedicated caretaker ensures consistent cleaning and maintenance of sanitation facilities. This regular upkeep reduces the risk of costly repairs and extends the lifespan of the infrastructure, contributing to the system's long-term sustainability (Brikké, 2000)
repairperson	Indicates whether a skilled individual can repair the sanitation facility when needed.	no yes	Ensuring that sufficient expertise is available within or near the community to undertake simple maintenance exercises can help keep facilities in good shape (Nelson et al., 2021).
desludgingneeded	Indicates whether desludging is required for the sanitation facility. It is determined by observing the facility's current condition, indicating if the waste has accumulated to a level that necessitates removal.	no yes	Emptying minimizes the risk that the fecal sludge and/or effluent overflows into the toilet and/or into water bodies or onto open ground (Conaway et al., 2023), and overfilled pits can collapse due to the weight of accumulated waste (Nakagiri et al., 2015)
emptiedonce	Indicates whether the sanitation facility has been emptied at once in the past	no yes	Inadequate emptying can lead to the restricted use of sanitation systems. When pits or tanks are not emptied regularly, they may become clogged or overloaded, reducing

			their ability to process waste effectively (Cookey et al., 2020; Thye et al., 2011)
structuraldamage	Indicates whether repairs are needed on the superstructure as indicated by evidence of cracking or damage to the toilet pedestal or squat slab.	no yes	Repairs to the superstructure ensure sanitation facility stability, usability, and hygiene by addressing cracks, damage, and structural weaknesses that could lead to service disruptions (Lüthi et al., 2011).
floorcondition	Assesses the condition of the floor or slab based on visible damage, wear, and tear.	severely damaged moderately damaged good condition	Maintain structural integrity and prevent further deterioration, as well as create a seamless surface that is much easier to clean and maintain (Brikké, 2000).
visibleexcreta	Indicates whether urine or feces are visible on the floor of the sanitation facility.	no yes	A facility with visible waste may indicate inadequate cleaning, poor drainage, or structural issues, which can lead to service disruptions and decreased usability. (Jenkins & Scott, 2007). Individuals may prefer open defecation, mainly when sanitation options are unappealing or unhygienically maintained (Dreibelbis et al., 2015).
cleansematerial	Assesses whether sanitation users use appropriate cleansing materials, such as toilet paper or water, at the facility.	not available available	The use of non-degradable cleansing materials, such as magazines, plastic bags, or stones, could clog the system, reduce the effective volume of the system, lead to rapid latrine filling, cause problems with emptying, and reduced treatment efficiency (Brikké & Bredero, 2003; Graham & Polizzotto, 2013; McMahon et al., 2011).
sharingfacility	Indicates whether a household uses their sanitation facility with multiple other households or members of the public.	no yes	Shared sanitation facilities frequently accommodate numerous users, raising the likelihood of blockages or backups. They are also more susceptible to physical issues like cracked pipes, broken sewer lines, and leaky joints due to regular wear and tear. Furthermore, the absence of accountability often results in neglected maintenance (Antwi-Agyei et al., 2020; Simiyu, 2016).
watersupply	Indicates whether there is continuous water supply throughout the year	continuous non-continuous	Access to adequate water supplies is essential for the operation, maintenance, and cleaning of sanitation systems and personal and domestic hygiene. In some cultures, water is also crucial for flushing, sewerage, and cleansing after defecation (Tilley et al., 2014).

[a] Short names as used in the Bayesian network analysis program.

Table S2. Factors influencing overflowing in 17,404 pit latrines across 12 Sub-Saharan African countries as identified through Bayesian Network analysis.

Node	Mutual info	Percent	Variance of beliefs
Overflowing	0.87909	100	0.2092790
Desludging needed	0.15721	17.9	0.0457274
Emptied frequently	0.00948	1.08	0.0030599
Structural damage	0.00313	0.356	0.0009236
Emptied at least once	0.00116	0.132	0.0003198
Flood risk	0.00057	0.0651	0.0001648
Visible excreta	0.00034	0.0385	0.0000984
Repair person	0.00004	0.00473	0.0000121
Flies present	0.00001	0.00167	0.0000043
Shared facility	0.00001	0.000698	0.0000018
Cleaning person	0.00000	0.000283	0.0000007
Cleansing material	0.00000	0.000153	0.0000004
Floor condition	0.00000	0.000122	0.0000003
Facility fee	0.00000	5.38e-05	0.0000001
Water supply	0.00000	0.000000	0.0000000

Table S3. Factors influencing evidence of overflowing in 1,130 septic systems across 12 Sub-Saharan African countries, as identified through Bayesian Network analysis.

Node	Mutual info	Percent	Variance of beliefs
Overflowing	0.65313	100	0.1397826
Desludging needed	0.09856	15.1	0.0274483
Flood risk	0.01060	1.62	0.0017110
Structural damage	0.00248	0.379	0.0004584
Emptied frequently	0.00215	0.329	0.0005630
Emptied at least once	0.00110	0.169	0.0002893
Visible excreta	0.00069	0.106	0.0001329
Facility fee	0.00011	0.0169	0.0000230
Flies present	0.00004	0.00605	0.0000076
Shared facility	0.00003	0.00482	0.0000061
Water supply	0.00001	0.00196	0.0000025
Cleansing material	0.00001	0.00103	0.0000013
Repair person	0.00000	0.000142	0.0000002
Cleaning person	0.00000	2.63e-05	0.0000000
Floor condition	0.00000	0.0000000	0.0000000

Fig S1. Output results for Model 1a: Outcome = rate of overflows; Sanitation type = pit latrines and septic systems.

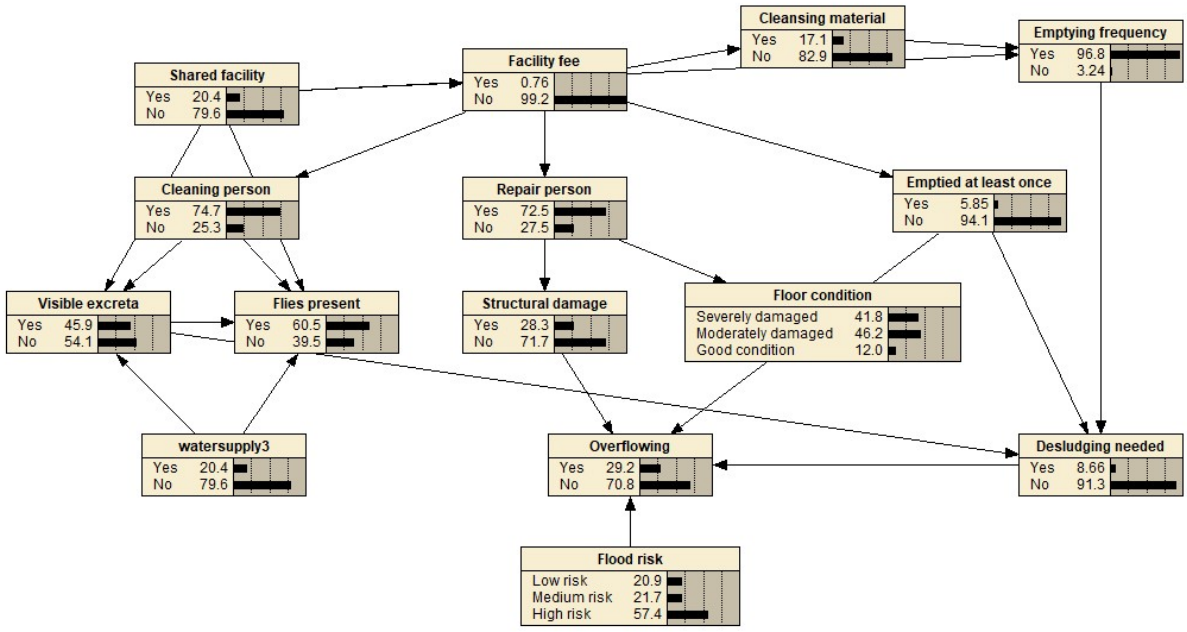

Table S4. Scoring Rule Results for Model 1a.

Logarithmic loss	0.4807
Quadratic loss	0.3109
Spherical payoff	0.8285
Gini coeff	0.3248
Area under ROC	0.6624

Table S5. Logistic regression results for variables included in Model 1a.

overflowing3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
visibleexcreta3	.978297	.0841813	-0.25	0.799	.8264675	1.158019
fliespresent3	1.35697	.1239038	3.34	0.001	1.134613	1.622903
floorcondition3	1.253678	.0927101	3.06	0.002	1.084524	1.449215
structuraldamage3	1.176526	.1227732	1.56	0.119	.9589071	.443531
desludgingneeded3	1	(omitted)				
emptiedonce3	.6127957	.2040023	-1.47	0.141	.3191149	1.17675
cleansematerial3	1.085078	.1200951	0.74	0.461	.8734758	1.347941
repairperson3	.8330187	.0899345	-1.69	0.091	.6741511	.029325
cleaningperson3	.9238839	.1034707	-0.71	0.480	.7417999	1.150663
facilityfee3	.4213924	.2605576	-1.40	0.162	.1254201	1.415814
sharingfacility3	.9728193	.1028486	-0.26	0.794	.7907538	1.196804
watersupply3	1.129232	.1163115	1.18	0.238	.9228034	1.381839
emptfreq3	2.432792	1.307306	1.65	0.098	.8485867	6.974513
floodrisk	1.098605	.0591491	1.75	0.081	.9885819	1.220873
_cons	.0818331	.0454931	-4.50	0.000	.0275248	.2432951

Fig S2. Output results for Model 1b: Outcome = rate of overflows; Sanitation type = pit latrines.

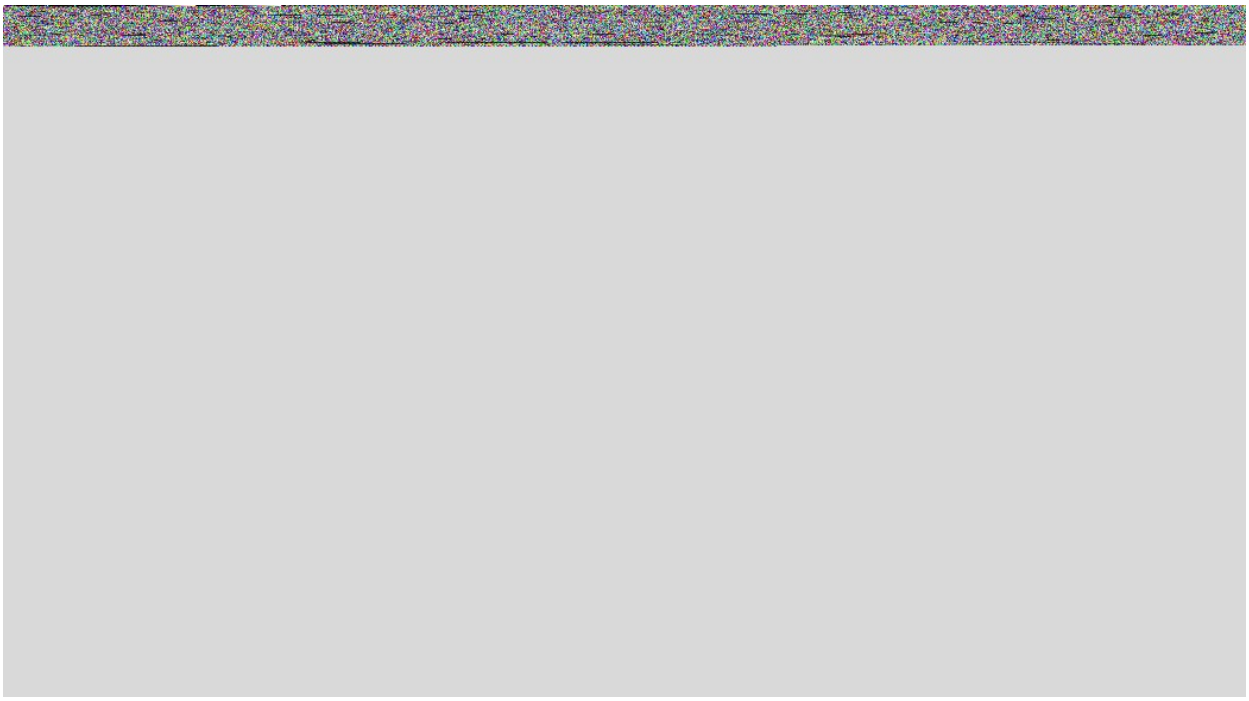

Table S6. Scoring Rule Results for Model 1b.

Logarithmic loss	0.4884
Quadratic loss	0.3180
Spherical payoff	0.8240
Gini coeff	0.3372
Area under ROC	0.6686

Table S7. Logistic regression results for variables included in Model 1b.

overflowing3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
visibleexcreta3	1.143116	.0445475	3.43	0.001	1.059055	1.233848
fliespresent3	1.347792	.0553108	7.27	0.000	1.243630	1.460678
floorcondition3	1.232766	.040337	6.40	0.000	1.156189	1.314416
structuraldamage3	1.170411	.0547461	3.36	0.001	1.067883	1.282784
desludgingneeded3	1	(omitted)				
emptiedonce3	.6867384	.0914618	-2.82	0.005	.5289637	.8915729
cleansematerial3	1.004017	.0528587	0.08	0.939	.9055819	1.113152
repairperson3	1.000978	.0489211	0.02	0.984	.9095437	1.101604
cleaningperson3	.7324072	.0374562	-6.09	0.000	.6625538	.8096254
facilityfee3	.5511781	.1473073	-2.23	0.026	.3264381	.9306432
sharingfacility3	1.203683	.0564382	3.95	0.000	1.097996	1.319541
watersupply3	1.196066	.0550211	3.89	0.000	1.092945	1.308916
emptfreq3	2.287205	.4776572	3.96	0.000	1.518942	3.444046
floodrisk	1.080804	.0260969	3.22	0.001	1.030846	1.133183
_cons	.0906427	.0199086	-10.93	0.000	.0589354	.139408

Fig S3. Output results for Model 1c: Outcome = rate of overflows; Sanitation type = septic systems.

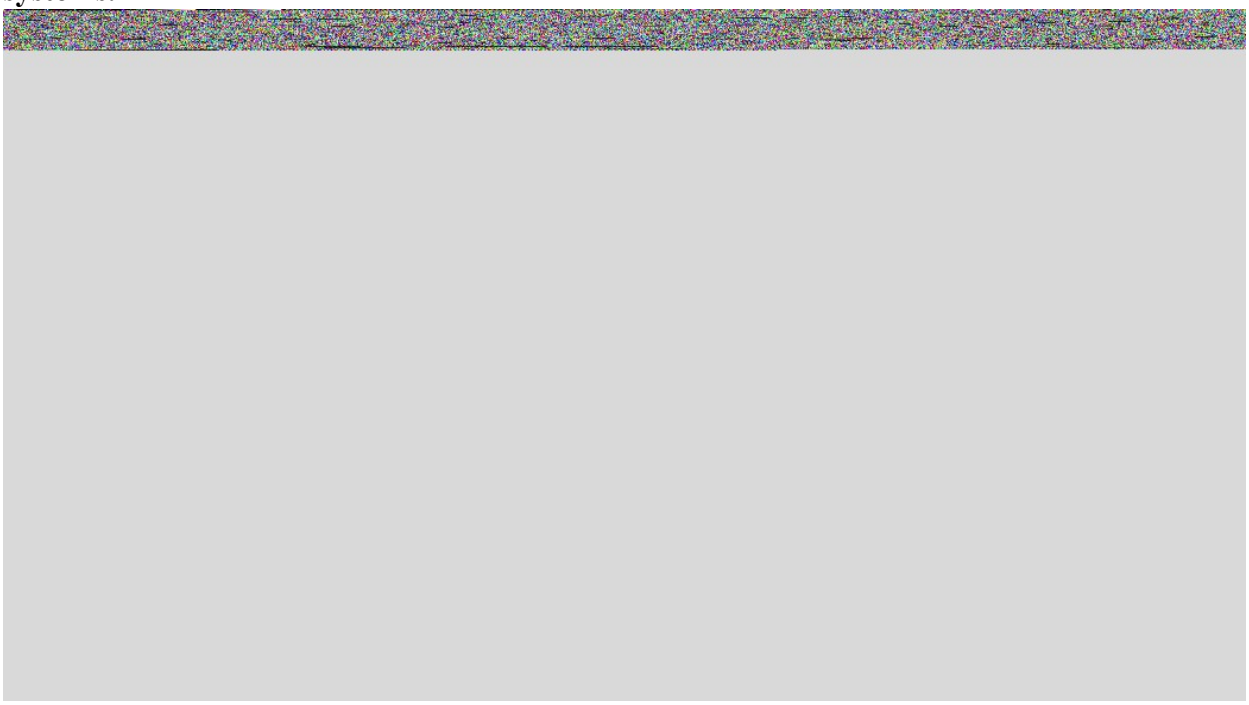

Table S8. Scoring Rule Results for Model 1c.

Logarithmic loss	0.3748
Quadratic loss	0.2267
Spherical payoff	0.8778
Gini coeff	0.4447
Area under ROC	0.7224

Table S9. Logistic regression results for variables included in Model 1c.

overflowing3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
visibleexcreta3	1.415712	.5415864	0.91	0.364	.668874	2.99644
fliespresent3	1.575058	.6739984	1.06	0.288	.6808491	3.643699
floorcondition3	2.546186	.8758004	2.72	0.007	1.297496	4.996596
structuraldamage3	.3831796	.1964728	-1.87	0.061	.1402666	1.046768
desludgingneeded3	1	(omitted)				
emptiedonce3	1	(omitted)				
cleansematerial3	4.179092	1.496209	3.99	0.000	2.071728	8.430072
repairperson3	1.050551	.5379795	0.10	0.923	.3850555	2.866231
cleaningperson3	.9111961	.3854782	-0.22	0.826	.397662	2.0879
facilityfee3	1	(omitted)				
sharingfacility3	.5774304	.2477484	-1.28	0.201	.2490513	1.338784
watersupply3	1.788926	.8811268	1.18	0.238	.6812982	4.697292
emptfreq3	1	(omitted)				
floodrisk	2.426986	.8309772	2.59	0.010	1.240577	4.748
_cons	.008504	.0075937	-5.34	0.000	.0014775	.0489452

Fig S4. Output results for Model 2a: Outcome = rate of overflows; Sanitation type = pit latrines and septic systems.

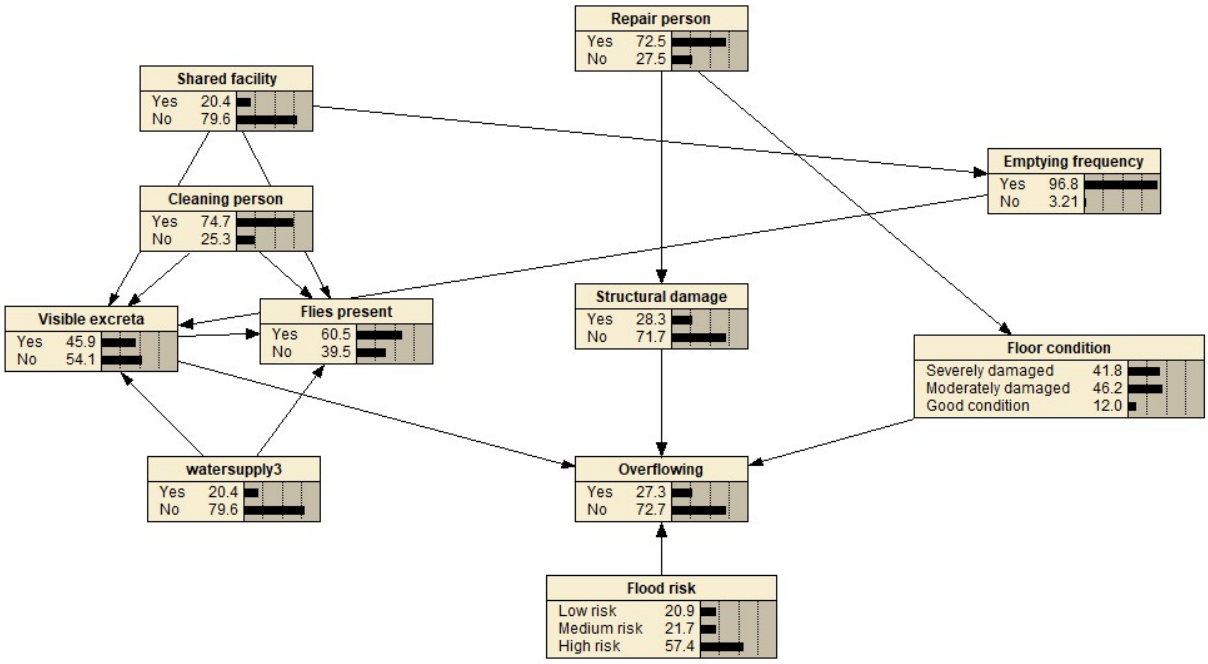

Table S10. Scoring Rule Results for Model 2a.

Logarithmic loss	0.5671
Quadratic loss	0.3811
Spherical payoff	0.7856
Gini coeff	0.2579
Area under ROC	0.6289

Table S11. Logistic regression results for variables included in Model 2a.

overflowing3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
visibleexcreta3	1.040961	.0813339	0.51	0.607	.8931551	1.213226
fliespresent3	1.337829	.1109207	3.51	0.000	1.137173	1.57389
floorcondition3	1.401595	.0930124	5.09	0.000	1.230652	1.596283
structuraldamage3	1.671065	.1515854	5.66	0.000	1.398876	1.996215
repairperson3	.8315463	.0813717	-1.89	0.059	.6864223	1.007353
cleaningperson3	.7910916	.0789523	-2.35	0.019	.6505421	.9620067
sharingfacility3	.9840871	.0926037	-0.17	0.865	.8183415	1.183403
watersupply3	1.075824	.1014013	0.78	0.438	.8943580	1.294109
emptfreq3	1.927976	.5305106	2.39	0.017	1.124298	3.306147
floodrisk	1.006762	.0485239	0.14	0.889	.9160113	1.106504
_cons	.1375702	.0419852	-6.50	0.000	.0756389	.2502093

Fig S5. Output results for Model 2b: Outcome = rate of overflows; Sanitation type = pit latrines.

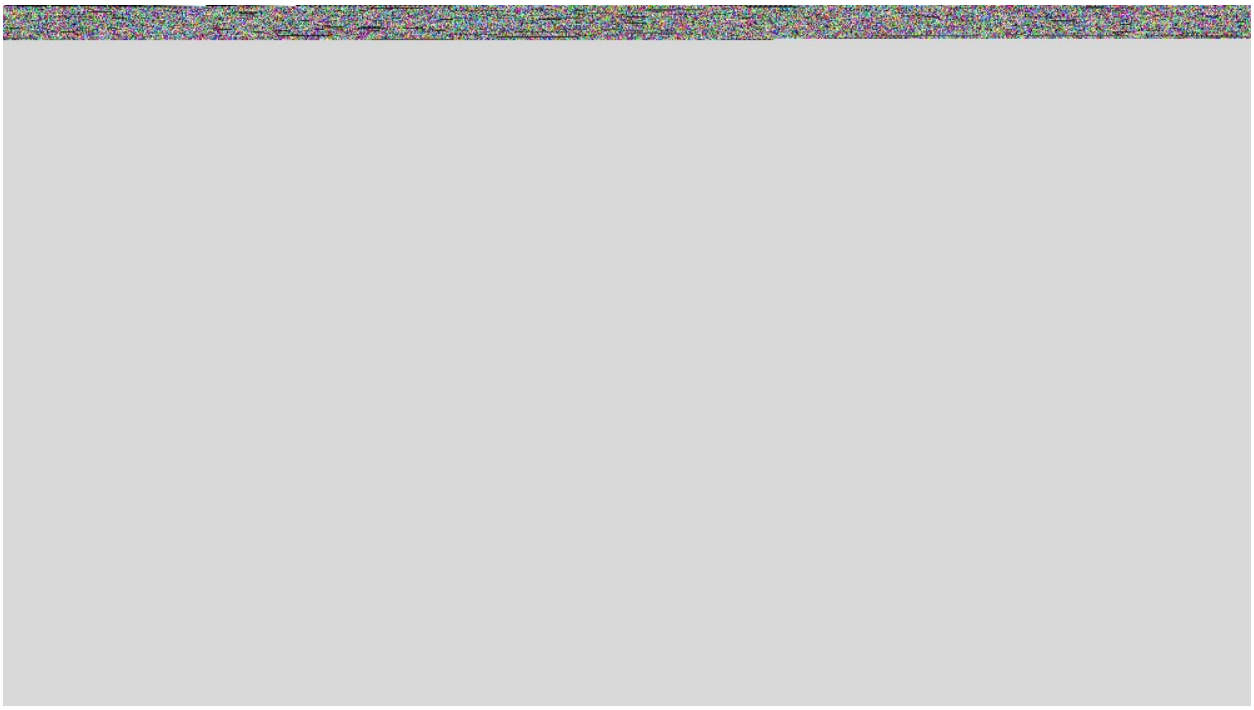

Table S12. Scoring Rule Results for Model 2b.

Logarithmic loss	0.5685
Quadratic loss	0.3821
Spherical payoff	0.7850
Gini coeff	0.2260
Area under ROC	0.6130

Table S13. Logistic regression results for variables included in Model 2b.

overflowing3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
visibleexcreta3	1.156722	.0409638	4.11	0.000	1.079157	1.239861
fliespresent3	1.4104	.0529882	9.15	0.000	1.310276	1.518174
floorcondition3	1.38881	.0409645	11.14	0.000	1.310798	1.471465
structuraldamage3	1.623311	.0664474	11.84	0.000	1.498163	1.758912
repairperson3	.9618331	.0427235	-0.88	0.381	.8816381	1.049323
cleaningperson3	.6810056	.0313041	-8.36	0.000	.6223335	.7452092
sharingfacility3	1.166604	.049334	3.64	0.000	1.073809	1.267417
watersupply3	1.196699	.0501928	4.28	0.000	1.102258	1.299232
emptfreq3	1.520031	.1600559	3.98	0.000	1.236582	1.868453
floodrisk	1.03312	.0224587	1.50	0.134	.9900266	1.07809
_cons	.1612441	.019568	-15.04	0.000	.1271116	.2045421

Fig S6. Output results for Model 2c: Outcome = rate of overflows; Sanitation type = septic systems.

Table S14. Scoring Rule Results for Model 2c.

Logarithmic loss	0.4537
Quadratic loss	0.2881
Spherical payoff	0.8417
Gini coeff	0.2874
Area under ROC	0.6437

Table S15. Logistic regression results for variables included in Model 2c.

overflowing3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
visibleexcreta3	1.350707	.2541291	1.60	0.110	.9341396	1.953038
fliespresent3	1.046466	.2079297	0.23	0.819	.7089147	1.544744
floorcondition3	2.331651	.3768038	5.24	0.000	1.698659	3.200523
structuraldamage3	.7641346	.1587888	-1.29	0.195	.508497	1.148289
repairperson3	.8394267	.2064261	-0.71	0.477	.5183954	1.359266
cleaningperson3	1.312417	.2633633	1.35	0.175	.8856471	1.944835
sharingfacility3	1.001603	.1897403	0.01	0.993	.6909507	1.451926
watersupply3	1.014786	.2481948	0.06	0.952	.628332	1.638929
emptfreq3	1	(omitted)				
floodrisk	1.730712	.2401616	3.95	0.000	1.318585	2.27165
_cons	.0407979	.0155253	-8.41	0.000	.0193518	.0860113

Fig S7. Output results for Model 3a: Outcome = combined rate of overflows and leakages; Sanitation type = pit latrines and septic systems.

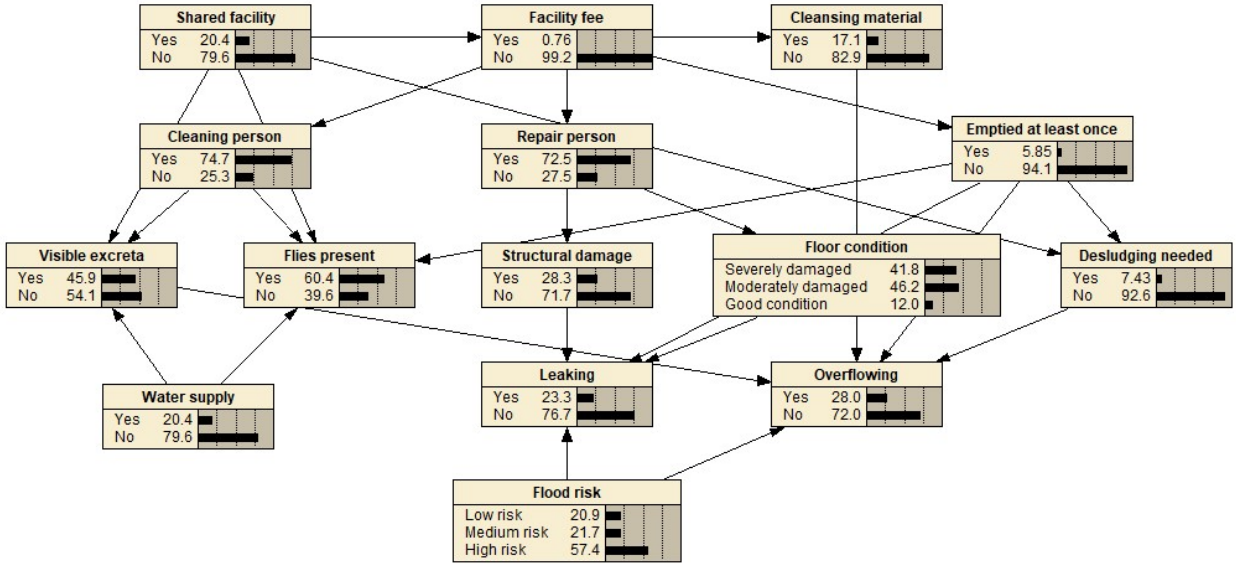

Table S16. Scoring Rule Results for Model 3a.

	Overflows	**Leakages**
Error rate	19.94%	22.61%
Scoring rule results		
Logarithmic loss	0.483	0.5221
Quadratic loss	0.3124	0.3406
Spherical payoff	0.8277	0.8112
Gini coeff	0.3077	0.1864
Area under ROC	0.6539	0.5932

Table S17. Logistic regression results for variables included in Model 3a (outcome = overflows).

overflowing3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
floodrisk	1.085584	.025765	3.46	0.001	1.036242	1.137275
visibleexcreta3	1.110549	.0422535	2.76	0.006	1.030746	1.19653
fliespresent3	1.336559	.0538252	7.20	0.000	1.23512	1.44633
structuraldamage3	1.126676	.0516701	2.60	0.009	1.029822	1.232638
floorcondition3	1.277343	.0410364	7.62	0.000	1.199393	1.360359
desludgingneeded3	1	(omitted)				
emptiedonce3	.4849606	.0500188	-7.02	0.000	.3961993	.5936073
cleansematerial3	1.057019	.053023	1.11	0.269	.958041	1.166222
repairperson3	.956218	.0455583	-0.94	0.347	.8709676	1.049813
cleaningperson3	.7715244	.0382094	-5.24	0.000	.7001551	.8501686
facilityfee3	.5503665	.146493	-2.24	0.025	.3266515	.927298
sharingfacility3	1.173796	.0536405	3.51	0.000	1.073233	1.283781
watersupply3	1.218332	.0549345	4.38	0.000	1.115283	1.330903
_cons	.1961189	.0130092	-24.56	0.000	.1722092	.2233482

Table S18. Logistic regression results for variables included in Model 3a (outcome = leakage).

leaking3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
floodrisk	1.081079	.0243549	3.46	0.001	1.034382	1.129883
visibleexcreta3	1.132832	.0410617	3.44	0.001	1.055145	1.216239
fliespresent3	1.325269	.0513402	7.27	0.000	1.228369	1.429812
structuraldamage3	1.126459	.049003	2.74	0.006	1.034395	1.226716
floorcondition3	1.297117	.0395106	8.54	0.000	1.221944	1.376915
desludgingneeded3	2.208926	.1372326	12.76	0.000	1.955685	2.494958
emptiedonce3	.3809119	.0376363	-9.77	0.000	.3138492	.4623045
cleaningperson3	.7572879	.0307268	-6.85	0.000	.6993968	.8199707
facilityfee3	.703359	.1592601	-1.55	0.120	.451275	1.096258
sharingfacility3	1.157087	.0504578	3.35	0.001	1.0623	1.260332
watersupply3	1.197823	.0514412	4.20	0.000	1.101127	1.303011
_cons	.1925941	.0118267	-26.82	0.000	.1707549	.2172265

Fig S8. Output results for Model 3b: Outcome = combined rate of overflows and leakages; Sanitation type = pit latrines.

Table S19. Scoring Rule Results for Model 3b.

	Overflowing	**Leaking**
Error rate	72.3%	23.44%
Scoring Rule Results		
Logarithmic loss	0.6931	0.5315
Quadratic loss	0.5	0.349
Spherical payoff	0.7071	0.806
Gini coeff	0	0.2059
Area under ROC	0.5	0.603

Table S20. Logistic regression results for variables included in Model 3b (outcome = overflows).

overflowing3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
floodrisk	1.081389	.026102	3.24	0.001	1.031422	1.133778
visibleexcreta3	1.144008	.0445614	3.45	0.001	1.05992	1.234767
fliespresent3	1.346963	.055264	7.26	0.000	1.242888	1.459752
structuraldamage3	1.173166	.0548552	3.42	0.001	1.070431	1.28576
floorcondition3	1.229172	.0401989	6.31	0.000	1.152856	1.310541
desludgingneeded3	1	(omitted)				
emptiedonce3	.4647968	.048011	-7.42	0.000	.3796109	.5690986
cleansematerial3	1.008558	.0530764	0.16	0.871	.9097154	1.11814
repairperson3	1.000504	.0488943	0.01	0.992	.9091192	1.101075
cleaningperson3	.7319553	.0374279	-6.10	0.000	.6621541	.8091146
facilityfee3	.5669939	.1516089	-2.12	0.034	.3357183	.9575948
sharingfacility3	1.200696	.0562555	3.90	0.000	1.095349	1.316176
watersupply3	1.199823	.055177	3.96	0.000	1.096409	1.312991
_cons	.2074241	.0139769	-23.34	0.000	.1817617	.2367096

Table S21. Logistic regression results for variables included in Model 3b (outcome = leakages).

leaking3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
floodrisk	1.076717	.0247537	3.22	0.001	1.029277	1.126343
visibleexcreta3	1.163909	.0431894	4.09	0.000	1.082265	1.251713
fliespresent3	1.341328	.0529945	7.43	0.000	1.24138	1.449323
structuraldamage3	1.171125	.0519829	3.56	0.000	1.073547	1.277573
floorcondition3	1.254914	.0390172	7.30	0.000	1.180725	1.333765
desludgingneeded3	2.132668	.1355589	11.92	0.000	1.882861	2.415618
emptiedonce3	.3684645	.0364705	-10.09	0.000	.3034898	.4473498
cleansematerial3	1.00197	.0511586	0.04	0.969	.9065549	1.107428
repairperson3	.9634287	.0447522	-0.80	0.423	.8795903	1.055258
cleaningperson3	.7476562	.0362655	-6.00	0.000	.6798512	.8222238
facilityfee3	.7044116	.1636016	-1.51	0.131	.4468198	1.110505
sharingfacility3	1.176435	.0527475	3.62	0.000	1.077464	1.284497
watersupply3	1.180124	.0516645	3.78	0.000	1.083087	1.285856
_cons	.2072787	.0133481	-24.44	0.000	.1827006	.2351631

Fig S9. Output results for Model 3c: Outcome = combined rate of overflows and leakages; Sanitation type = septic systems.

Table S22. Scoring Rule Results for Model 3c.

	Overflowing	Leaking
Error rate	81.13	14.82
Scoring Rule Results		
Logarithmic loss	0.6931	0.3838
Quadratic loss	0.5	0.2349
Spherical payoff	0.7071	0.872
Gini coeff	0	0.4094
Area under ROC	0.5	0.7047

Table S23. Logistic regression results for variables included in Model 3c (outcome = overflows).

overflowing3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
floodrisk	2.253313	.420254	4.36	0.000	1.563393	3.247692
visibleexcreta3	1.226525	.2594316	0.97	0.334	.810275	1.856609
fliespresent3	1.193657	.2722062	0.78	0.438	.7634282	1.86634
structuraldamage3	.3136585	.0919235	-3.96	0.000	.1766024	.5570801
floorcondition3	2.192711	.4055738	4.24	0.000	1.525947	3.150819
desludgingneeded3	1	(omitted)				
emptiedonce3	1	(omitted)				
cleaningperson3	1.574651	.3637389	1.97	0.049	1.00129	2.476332
facilityfee3	1	(omitted)				
sharingfacility3	.8172662	.1885244	-0.87	0.382	.5200093	1.284446
watersupply3	1.145086	.3072353	0.50	0.614	.6767907	1.937412
_cons	.0167654	.0076903	-8.91	0.000	.0068229	.0411964

Table S24. Logistic regression results for variables included in Model 3c (outcome = leakages).

leaking3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
floodrisk	2.059731	.3452181	4.31	0.000	1.483017	2.860716
visibleexcreta3	1.268503	.2592575	1.16	0.245	.8498112	1.89348
fliespresent3	1.036811	.2231064	0.17	0.867	.680039	1.580759
structuraldamage3	.3533482	.0938181	-3.92	0.000	.2099901	.5945752
floorcondition3	2.123009	.3693778	4.33	0.000	1.509569	2.98573
desludgingneeded3	5.057234	1.65032	4.97	0.000	2.667729	9.587039
emptiedonce3	1	(omitted)				
cleaningperson3	1.547284	.3361394	2.01	0.045	1.010767	2.368586
facilityfee3	.8283498	.9617094	-0.16	0.871	.0851086	8.062207
sharingfacility3	.8966685	.1912449	-0.51	0.609	.5903159	1.362007
watersupply3	1.026012	.2707138	0.10	0.922	.6117342	1.720848
_cons	.0214733	.0089723	-9.19	0.000	.0094675	.0487035

Fig S10. Output results for Model 4a: Outcome = leakages; Sanitation type = pit latrines and septic systems.

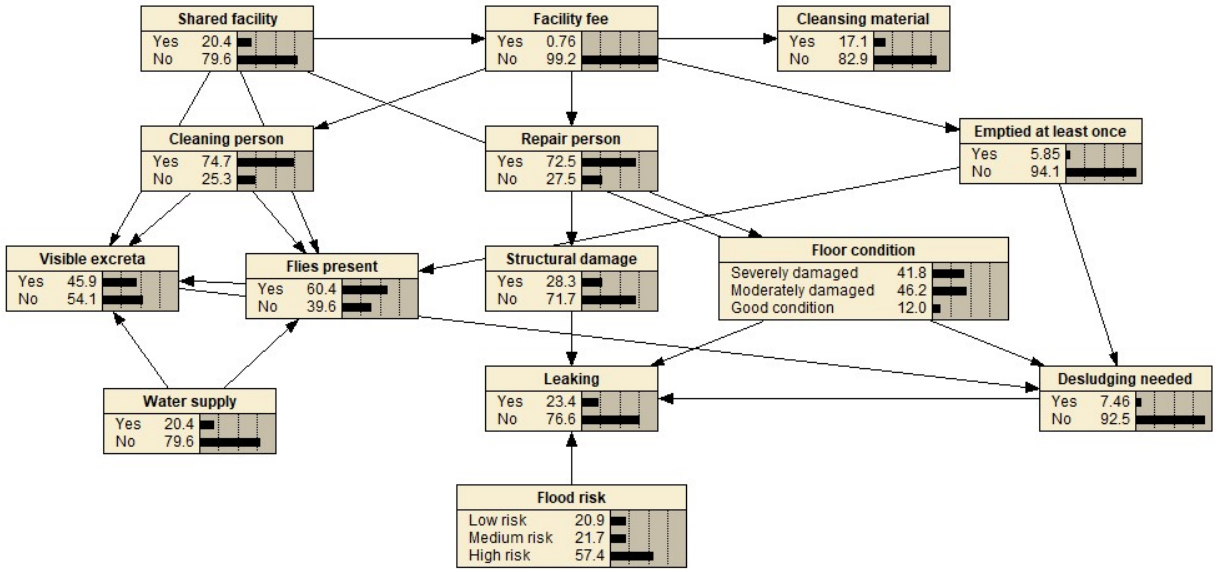

Table S25. Scoring Rule Results for Model 4a.

	leaking
Error rate	22.31
Scoring Rule Results	
Logarithmic loss	0.5225
Quadratic loss	0.3405
Spherical payoff	0.8114
Gini coeff	0.1722
Area under ROC	0.5861

Table S26. Logistic regression results for variables included in Model 4a.

leaking3	**Odds ratio**	**Std. err.**	**z**	**P>z**	**[95% conf.**	**interv**
floodrisk	1.079952	.0244001	3.40	0.001	1.033172	1.12885
visibleexcreta3	1.132935	.0410705	3.44	0.001	1.055232	1.216361
fliespresent3	1.327276	.0514424	7.31	0.000	1.230185	1.432029
structuraldamage3	1.126583	.0490427	2.74	0.006	1.034448	1.226925
floorcondition3	1.298691	.0396552	8.56	0.000	1.223248	1.378787
desludgingneeded3	2.204092	.1370288	12.71	0.000	1.951239	2.489712
emptiedonce3	.3823478	.0377993	-9.73	0.000	.314998	.4640976
repairperson3	.9276641	.042006	-1.66	0.097	.8488816	1.013758
cleaningperson3	.7840607	.0368521	-5.18	0.000	.715059	.859721
sharingfacility3	1.152709	.0503611	3.25	0.001	1.058111	1.255764
facilityfee3	.6971687	.1584768	-1.59	0.113	.4465263	1.0885
cleansematerial3	1.048379	.0510081	0.97	0.332	.9530233	1.153274
watersupply3	1.194647	.0513419	4.14	0.000	1.098141	.299635
_cons	.1967324	.0124653	-25.66	0.000	.173757	.2227456

Fig S11. Output results for Model 4b: Outcome = leakages; Sanitation type = pit latrines.

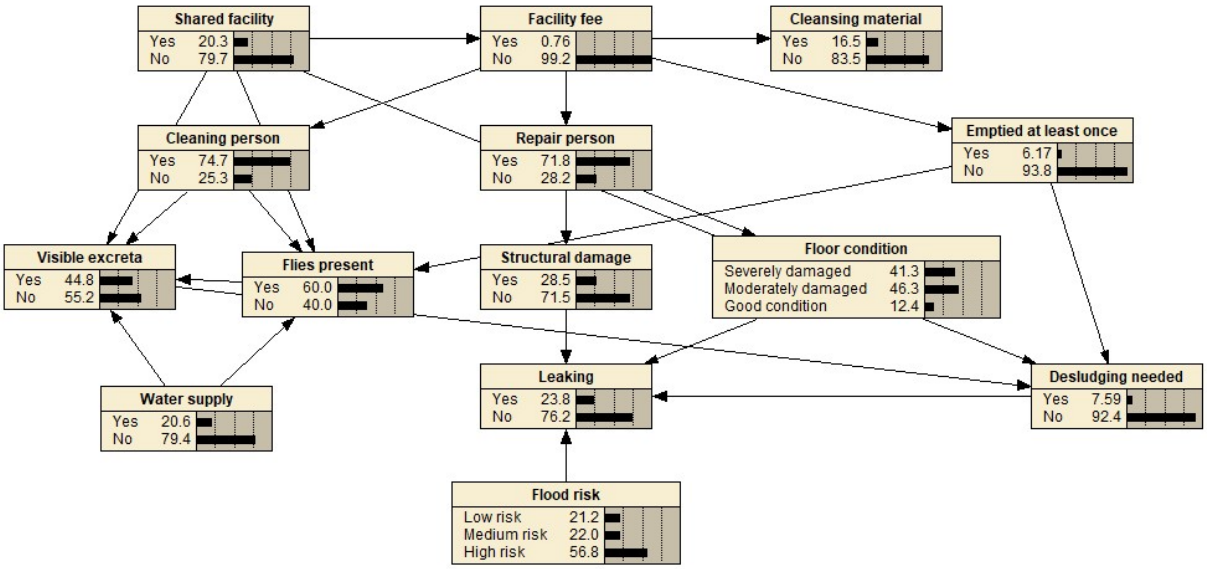

Table S27. Scoring Rule Results for Model 4b.

	leaking	
Error rate	23.02%	
Scoring Rule Results		
Logarithmic loss	0.5324	
Quadratic loss	0.3491	
Spherical payoff	0.8061	
Gini coeff	0.164	
Area under ROC	0.582	(ROC curve dips below main diagonal)

Table S28. Logistic regression results for variables included in Model 4b.

leaking3	Odds ratio	Std. err.	z	P>z	[95% conf.	interval]
floodrisk	1.076717	.0247537	3.22	0.001	1.029277	1.126343
visibleexcreta3	1.163909	.0431894	4.09	0.000	1.082265	1.251713
fliespresent3	1.341328	.0529945	7.43	0.000	1.24138	1.449323
structuraldamage3	1.171125	.0519829	3.56	0.000	1.073547	1.277573
floorcondition3	1.254914	.0390172	7.30	0.000	1.180725	1.333765
desludgingneeded3	2.132668	.1355589	11.92	0.000	1.882861	2.415618
emptiedonce3	.3684645	.0364705	-10.09	0.000	.3034898	.4473498
repairperson3	.9634287	.0447522	-0.80	0.423	.8795903	1.055258
cleaningperson3	.7476562	.0362655	-6.00	0.000	.6798512	.8222238
sharingfacility3	1.176435	.0527475	3.62	0.000	1.077464	1.284497
facilityfee3	.7044116	.1636016	-1.51	0.131	.4468198	1.110505
cleansematerial3	1.00197	.0511586	0.04	0.969	.9065549	1.107428
watersupply3	1.180124	.0516645	3.78	0.000	1.083087	1.285856
_cons	.2072787	.0133481	-24.44	0.000	.1827006	.2351631

Fig S12. Output results for Model 4c: Outcome = leakages; Sanitation type = septic systems

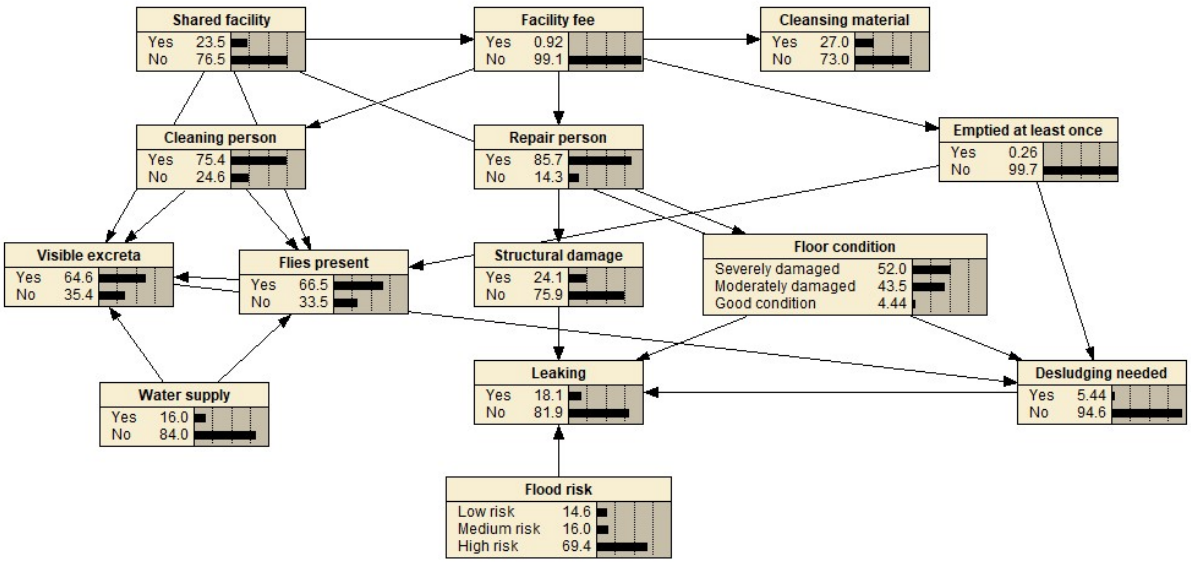

Table S29. Scoring Rule Results for Model 4c.

	leaking
Error rate	16.17
Scoring Rule Results	
Logarithmic loss	0.3924
Quadratic loss	0.244
Spherical payoff	0.866
Gini coeff	0.433
Area under ROC	0.7165

Table S30. Logistic regression results for variables included in Model 4c.

leaking3	**Odds ratio**	**Std. err.**	**z**	**P>z**	**[95% conf.**	**interval]**
floodrisk	1.98439	.3373869	4.03	0.000	1.422018	2.769164
visibleexcreta3	1.45764	.3098413	1.77	0.076	.9609835	2.21098
fliespresent3	1.141448	.2546722	0.59	0.553	.7371256	1.767545
structuraldamage3	.3539402	.0953432	-3.86	0.000	.2087546	.6001002
floorcondition3	2.154448	.3865374	4.28	0.000	1.515719	3.062339
desludgingneeded3	5.573121	1.864726	5.13	0.000	2.892628	10.73753
emptiedonce3	1	(omitted)				
repairperson3	1.461152	.4313811	1.28	0.199	.8192045	2.606146
cleaningperson3	1.20162	.2817002	0.78	0.433	.7589568	1.902469
sharingfacility3	.8826395	.1925656	-0.57	0.567	.5755414	1.353599
facilityfee3	.640687	.7748764	-0.37	0.713	.0598637	6.856908
cleansematerial3	2.892491	.556772	5.52	0.000	1.983468	4.218118
watersupply3	1.09578	.2954404	0.34	0.734	.6459895	1.858752
_cons	.0113863	.0054907	-9.28	0.000	.0044251	.029298

Fig S13. Sample design for the household survey