SUPPLEMENTARY MATERIAL

| Contents | Page |
|--|------|
| Supplementary methods | 2 |
| Figure S1 Cohort generation for (A) development study; (B) validation study for calculating relative risk (RR) and infection rate (IR); and (C) Cross-validating RR and IR | 3 |
| Figure S2 CONSORT diagram of the validation analysis in Trusted Research Environment for England | 4 |
| Figure S3 Baseline 1-year mortality in England (age \geq 30) by number of underlying conditions, age category, and sex in development (n=3,862,012 scaled up to mid-2018 population of England of age \geq 30) and validation cohorts (n=35,098,810) | 5 |
| Figure S4 Excess COVID-19 and total deaths over 1 year in England using Lancet 2020 model: (a) CPRD (Predicted), b) NHS Digital TRE (Predicted), and c) NHS Digital TRE (Observed) | 6 |
| Table S1 Cross-validated age-specific and overall infection rate for COVID-19 | 7 |
| Table S2 Cross-validated relative risk across two non-overlapping random sub- samples· | 7 |
| Table S3 Sensitivity analysis and 2-fold cross-validation of relative risk $({\bf RR})$ and population infection rate $({\bf IR})$ | 8 |
| Table S4 Assessment of risk ratio, rate ratio, infection risk, and infection rate of different periods based on vaccination in people of age 30 or older | 9 |
| Table S5 Baseline one year mortality risk in England per underlying condition (NHS Digital TRE; n = 35,098,810 age ≥30 years) | 10 |
| Table S6 Overlap between high-risk groups for CVD, DM, CKD, and COPD in England (NHS Digital TRE; n = 35,098,810 age ≥30) | 11 |

Supplementary methods

To address potential double-counting of individuals surviving to 1 March 2020 in both cohorts, we randomly divided the population into two sub-samples without replacement and cross-validated analysis of RR and IR by selecting the pre-pandemic unexposed group from the first sub-sample and COVID-19 exposed group from the second sub-sample, and vice-versa (**Figure S1.C**), and averaging results (**Table S1, S2**).

We used total contributed time of each patient within 1 year (i.e. time from the start of a period to either death or end of the period) in the survival analysis.

Sensitivity analysis

In sensitivity analyses of KM estimates, stratified by combinations of age, sex, and number of underlying conditions, the best performing model in terms of the estimated versus actual numbers of COVID-19 related deaths was the KM analysis stratified by all three explanatory variables of age, sex, and number of underlying conditions. To internally validate the model against overfitting or underfitting in cross-validated 50% sub-samples (Figure S1.C), we calculated RR and IR on different data fractions (training set) and applied the model on KM results on remaining data (validation set) (**Table S3**).

COVID-19 vaccination started in England in December 2020. To assess vaccination effects on overall RR and IR estimation, we divided the study period into quarters where the 4th quarter (December 2020 to March 2021) included those with 0, 1, or 2 doses of vaccination. We compared RR and IR values of the 4th quarter per vaccination dose to the corresponding quarter in pre-pandemic group (**Table S4**).

Figure S1 Cohort generation for (A) development study; (B) validation study for calculating relative risk (RR) and infection rate (IR); and (C) Cross-validating RR and IR

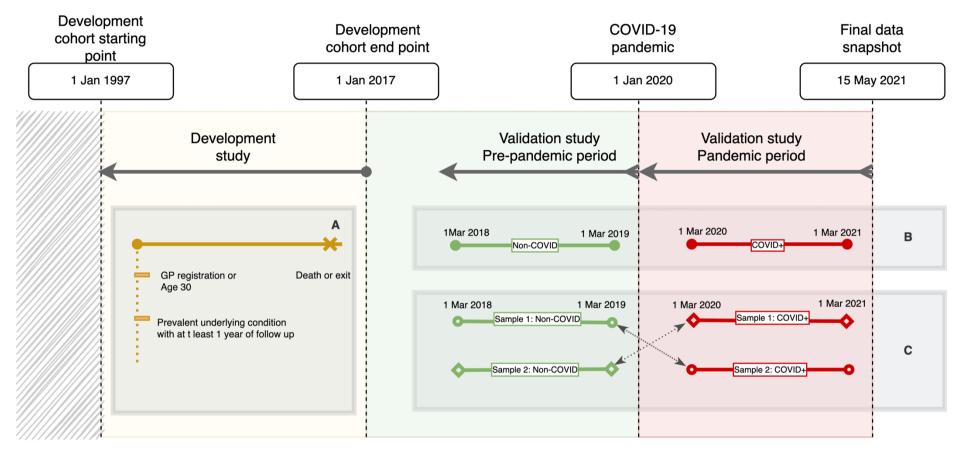
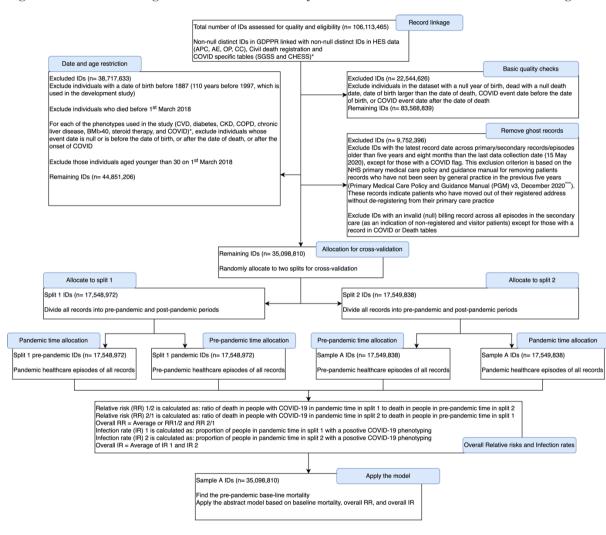


Figure S2 CONSORT diagram of the validation analysis in Trusted Research Environment for England



* For more details, refer to the "Data sources" sub-section of methods in the main manuscript-

** For more details, refer to the "Exposures and outcomes of interest" sub-section of methods in the main manuscript-

*** Primary Medical Care Policy and Guidance Manual (PGM) v3.26

Figure S3 Baseline one year mortality in England (age \ge 30) by number of underlying conditions, age category, and sex in development (n=3,862,012 scaled up to mid-2018 population of England of age \ge 30) and validation cohorts (n=35,098,810)

a) Development cohort (n=3,862,012) scaled up to the mid-2018 population of England aged 30 and over^{*}

| | | | | M | en | | | | | | | Wo | men | | | |
|--------------------------|-----------------------------|---------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|----------------------------|----------------------------|---------------------------|--------------------------|--------------------------|--------------------------|---------------------------|---------------------------|----------------------------|
| 3+ · suo | 4.6% 5281 (230) | 2.8% 4062 (111) | 7% 6574 (450) | 8.9% 9205 (798) | 12.2% 11158 (1302) | 15% 11745 (1706) | 19.4% 10076 (1880) | 28.7% 8765 (2412) | 4% 6455 (239) | 6.5% 4236 (266) | 5.2% 6079 (303) | 8.1% 8087 (633) | 8.4% 10168 (816) | 10.7% 11928 (1238) | 16.4% 10920 (1724) | 26% 13065 (3228) |
| of underlying conditions | 1.5% 42935 (624) | 3.5% 23966 (798) | 3.5% 32190 (1082) | 4.9% 38882 (1825) | 5.9% 41789 (2384) | 9% 42073 (3695) | 14.6% 29659 (4218) | 24% 25992 (5960) | 1.2% 52139 (615) | 1.1% 19904 (202) | 2.2% 26203 (551) | 3.1% 32217 (972) | 4.7% 37920 (1752) | 6.7% 44182 (2888) | 11.2% 39066 (4227) | 20.4% 53781 (10498) |
| Number of under 1. | 0.6% 494143 (2843) | 1.2% 143986 (1678) | 1.8% 159645 (2751) | 2.4% 157591 (3750) | 3.4% 153337 (5153) | 5.5% 133781 (7225) | 10.1% 86722 (8518) | 19.1% 72309 (13230) | 0.4% 568964 (2201) | 0.9% 122019 (1073) | 1% 133048 (1357) | 1.7% 138365 (2348) | 2.4% 151980 (3576) | 3.8% 165274 (6134) | 7% 139016 (9517) | 17.1% 182785 (30173) |
| UN 0. | 0.1% 11855088 (13964) | 0.4% 1117849 (4071) | 0.6% 837158 (5208) | 1.1% 627227 (6556) | 1.7% 501551 (8408) | 3% 389012 (11451) | 6% 215552 (12597) | 12.4% 167483 (20225) | 0.1% 11173934 (7867) | 0.2% 1117546 (2705) | 0.4% 876645 (3723) | 0.7% 699902 (4484) | 1.1% 624247 (6840) | 2.1% 580360 (12185) | 4.5% 410612 (17933) | 12% 489513 (56815) |
| | 30-55 | 56-60 | 61-65 | 66-70 | 71-75 | 76-80 | 81-85 | >85 Age grou | 30-55 up (years) | 56-60 | 61-65 | 66-70 | 71-75 | 76-80 | 81-85 | >85 |
| | | | | Me | n | | | | | | | Wor | nen | | | |

b) Validation cohort (n=35,095,810)*

2.4% 2.8% 3.3% 4.1% 4.9% 6.3% 8.6% 13.8% 1.6% 2% 2.6% 3.1% 3.9% 5% 6.5% 11.3% 3+ 22821 22178 34592 52294 72820 73340 65967 53672 25501 20450 29624 42653 59707 66354 66313 71987 (542) (627) (1151) (2152) (3574) (4656) (5648) (7414) (419) (419) (776) (1332) (2317) (3308) (4329) (8128) S lying conditio 15% 1.2% 1.6% 6% 12.1% 1.2% 1.5% 2.1% 2.8% 3.6% 5.3% 8% 0.8% 2.1% 2.9% 4.2% 124177 84027 106723 136565 164446 143820 123236 109709 134945 72967 88052 110540 137355 135527 132758 168146 (1482)(1300) (2203) (5927) (9846) (1039) (852) (1418) (2363) (3955) (7957) (20423) (3878) (7634)(16405) (5631) of under 0.6% 4.1% 6.9% 15.9% 0.4% 0.8% 1.1% 1.5% 2.1% 3% 5.4% 14.2% 1% 1.3% 1.8% 2.6% 814599 325465 335180 360503 374711 274455 205700 170808 921086 282066 284388 313291 345399 286165 246899 296360 (4830) (3153) (4397) (6541) (9648) (11367) (14255) (27151) (3703) (2256) (3074) (4590) (7101) (8641) (13234) (42166) Number 0.1% 0.4% 0.6% 1% 1.6% 2.9% 5.5% 14.7% 0.1% 0.3% 0.4% 0.6% 1% 2% 4.4% 14% 8182088 1256781 946287 782809 633105 346346 201717 136204 8798443 1340039 1062597 936253 801475 480510 313739 290076 0 (11496) (4838) (5726) (7577) (9927) (10094) (11080) (20037) (7305) (3548) (4151) (5699) (8269) (9839) (13706) (40467 30-55 30-55 56-60 61-65 66-70 71-75 76-80 81-85 >85 56-60 61-65 66-70 71-75 76-80 81-85 >85 Age group (years) 1-year all-cause deaths 0 20000 40000 60000 80000

* Each cell: Mortality risk %, number of people at risk, (number of deaths)

Figure S4 Excess COVID-19 and total deaths over 1 year in England using Lancet 2020 model: (a) CPRD (Predicted), b) NHS Digital TRE (Predicted), and c) NHS Digital TRE (Observed)

a) CPRD (Predicted)†: n=35,407,313 (scaled up from 3,862,012 to mid-2018 population of England aged 30 and over), non-COVID and indirect deaths= 356,186, excess deaths directly related to COVID= 74,628; total predicted number of deaths in England: 430,814^{*}

| 3+ · | 100 | 80 | 159 | 301 | 444 | 618 | 756 | 1183 |
|---------------------------------|---------|--------|---------|-------------------|---------------------|---------|---------|---------|
| | (569) | (457) | (912) | (1732) | (2562) | (3562) | (4360) | (6823) |
| rlying conditio | 260 | 211 | 343 | 587 | 867 | 1379 | 1770 | 3448 |
| | (1499) | (1211) | (1976) | (3384) | (5003) | (7962) | (10215) | (19906) |
| Number of underlying conditions | 1057 | 577 | 862 | 1278 | 1829 | 2799 | 3778 | 9090 |
| 1 | (6101) | (3328) | (4970) | (7376) | (10558) | (16158) | (21813) | (52493) |
| N | 4573 | 1420 | 1871 | 2313 | 3194 | 4951 | 6395 | 16135 |
| 0. | (26404) | (8196) | (10802) | (13353) | (18442) | (28587) | (36925) | (93175) |
| | 30-55 | 56-60 | 61-65 | 66-70 Age grou | 71-75 ip (years) | 76-80 | 81-85 | >85 |

c) NHS Digital TRE (Observed): n = 35,098,810, non-COVID and indirect deaths=458,393, excess deaths directly related to COVID= 127,020; total observed number of deaths in England: 585,413*

| | + - | 403 (1441) | 493 (1775) | 925 (3324) | 1649 (6105) | 2872 (10475) | 3940 (14555) | 4813 (18168) | 5679 (24309) |
|---------------------------------|-----|-----------------|-----------------|-----------------|-------------------|--------------------|-----------------|-----------------|------------------|
| rlying conditio | 2. | 739 (3412) | 807 (3269) | 1276 (5302) | 2063 (8533) | 3387 (14324) | 4771 (19158) | 6542 (25409) | 9993 (45561) |
| Number of underlying conditions | 1. | 1730 (9070) | 1354 (6259) | 1954 (8662) | 2810 (13050) | 4456 (20467) | 6100 (26287) | 8088 (34178) | 13635 (64223) |
| | 0- | 3444 (26299) | 2039 (12527) | 2400 (14564) | 3106 (18778) | 4435 (24962) | 5206 (25370) | 6028 (28115) | 9883 (47482) |
| | | 30-55 | 56-60 | 61-65 | 66-70 Age grou | 71-75 p (years) | 76-80 | 81-85 | >85 |
| | | | | 1 | year COVID-19 rel | ated excess death | s | | |
| | | | | 0 | 5000 10000 | 1500020000 | | | |

[†]Using observed infection rate over 1 year (IR: 6-27) and observed relative risk on 1-year mortality (RR: 4-34) from NHS Digital TRE ^{*}Each cell: Excess COVID-19 deaths, (Total number of deaths) TRE: Trusted Research Environment

b) NHS Digital TRE (Predicted)[†]: n = 35,098,810, non-COVID and indirect deaths=478,971, excess deaths directly related to COVID= 100,338; total predicted number of deaths in England: 579,309*

| 3+- | 202 | 220 | 405 | 730 | 1235 | 1669 | 2090 | 3256 |
|---------------------------------|---------|---------|---------|-------------------|--------------------|---------|---------|---------|
| Su | (1163) | (1266) | (2332) | (4214) | (7126) | (9633) | (12067) | (18798) |
| Number of underlying conditions | 529 | 452 | 759 | 1308 | 2071 | 2779 | 3729 | 7713 |
| | (3050) | (2604) | (4380) | (7549) | (11953) | (16044) | (21532) | (44541) |
| mber of unde | 1788 | 1134 | 1565 | 2332 | 3509 | 4191 | 5758 | 14517 |
| | (10321) | (6543) | (9036) | (13463) | (20258) | (24199) | (33247) | (83834) |
| N | 3938 | 1758 | 2070 | 2781 | 3811 | 4175 | 5192 | 12672 |
| 0- | (22739) | (10144) | (11947) | (16057) | (22007) | (24108) | (29978) | (73176) |
| | 30-55 | 56-60 | 61-65 | 66-70 Age grou | 71-75 p (years) | 76-80 | 81-85 | >85 |

| Infection rate* (incidence proportion) per 100 | | | | | | | | | | |
|--|---|---|--|---|--|--|--|--|--|--|
| | | | Age bands for | age-specific IR | | | | Overall | | |
| 30-55 | 56-60 | 61-65 | 66-70 | 71-75 | 76-80 | 81-85 | >85 | - | | |
| 7.54 | 6.36 | 5.16 | 3.59 | 3.10 | 3.69 | 5.07 | 8.73 | 6.27 | | |
| (7.524- | (6.337- | (5.138- | (3.569- | (3.080- | (3.664- | (5.033- | (8.681- | (6.264-6.280) | | |
| 7.549) | 6.378) | 5.188) | 3.614) | 3.122) | 3.718) | 5.105) | 8.776) | | | |
| 7.53 | 6.34 | 5.18 | 3.58 | 3.10 | 3.70 | 5.07 | 8.71 | 6.27 | | |
| (7.509- | (6.307- | (5.147- | (3.550- | (3.070- | (3.659- | (5.020- | (8.643- | (6.254-6.277) | | |
| 7.544) | 6.378) | 5.218) | 3.613) | 3.129) | 3.734) | 5.122) | 8.777) | | | |
| 7.55 | 6.38 | 5.14 | 3.60 | 3.10 | 3.69 | 5.07 | 8.75 | 6.28 | | |
| (7.530- | (6.346- | (5.109- | (3.571- | (3.074- | (3.647- | (5.015- | (8.680- | (6.268-6.291) | | |
| 7.565) | 6.417) | 5.179) | 3.634) | 3.132) | 3.723) | 5.117) | 8.815) | | | |
| 7.54 (7.520- | 6.36 (6.327- | 5.16 (5.128- | 3.59 (3.561- | 3.10 (3.072- | 3.69 (3.653- | 5.07 (5.018- | 8.73 (8.662- | 6.27 (6.261- | | |
| 7.554) | 6.397) | 5.199) | 3.623) | 3.131) | 3.728) | 5.120) | 8.796) | 6.284) | | |
| | 7.54 (7.524- 7.549) 7.53 (7.509- 7.544) 7.55 (7.530- 7.565) 7.54 (7.520- | $\begin{array}{ccccc} 7.54 & 6\cdot 36 \\ (7\cdot 524- & (6\cdot 337- \\ 7\cdot 549) & 6\cdot 378) \\ \hline 7\cdot 53 & 6\cdot 34 \\ (7\cdot 509- & (6\cdot 307- \\ 7\cdot 544) & 6\cdot 378) \\ \hline 7\cdot 55 & 6\cdot 38 \\ (7\cdot 530- & (6\cdot 346- \\ 7\cdot 565) & 6\cdot 417) \\ \hline 7\cdot 54 & (7\cdot 520- & 6\cdot 36 & (6\cdot 327- \\ \hline \end{array}$ | $\begin{array}{cccccccccccccccccccccccccccccccccccc$ | $\begin{tabular}{ c c c c c c c } \hline & Age bands for \\\hline\hline 30.55 & 56-60 & 61-65 & 66-70 \\\hline\hline 7.54 & 6.36 & 5.16 & 3.59 \\\hline (7.524- & (6\cdot337- & (5\cdot138- & (3\cdot569- \\ 7.549) & 6\cdot378) & 5\cdot188 & 3\cdot614) \\\hline\hline 7.53 & 6.34 & 5\cdot18 & 3.58 \\\hline (7.509- & (6\cdot307- & (5\cdot147- & (3\cdot550- \\ 7.544) & 6\cdot378) & 5\cdot218) & 3\cdot613) \\\hline 7.55 & 6\cdot38 & 5\cdot14 & 3\cdot60 \\\hline (7.530- & (6\cdot346- & (5\cdot109- & (3\cdot571- \\ 7.565) & 6\cdot417) & 5\cdot179) & 3\cdot634) \\\hline 7.54 & (7\cdot520- & 6\cdot36 & (6\cdot327- & 5\cdot16 & (5\cdot128- & 3\cdot59 & (3\cdot561- \\ \hline \end{tabular}$ | $\begin{array}{ c c c c c c c c c c c c c c c c c c c$ | $\begin{array}{ c c c c c c c c c c c c c c c c c c c$ | $\begin{array}{ c c c c c c c c c c c c c c c c c c c$ | $\begin{array}{ c c c c c c c c c c c c c c c c c c c$ | | |

Table S1 Cross-validated age-specific and overall infection rate for COVID-19

*The rate here does not denote the time in the denominator. The denominator is the total number of people at risk at the start of each period.

Table S2 Cross-validated relative risk across two non-overlapping random sub-samples

| | | | | Risk ratio | (relative risk) | | | | | | |
|----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|----------------|-----------------------------|--|--|
| | Age bands for age-specific | | | | | | | | | | |
| | 30-55 | 56-60 | 61-65 | 66-70 | 71-75 | 76-80 | 81-85 | >85 | - | | |
| Non-sampled relative risk | 2.35 | 3.41 | 4.65 | 6.79 | 7.89 | 7.36 | 5.65 | 2.99 | 4.35 | | |
| (95% CI) | $(2 \cdot 28 - 2 \cdot 42)$ | (3.30-3.53) | (4.52-4.79) | (6.63-6.95) | $(7 \cdot 75 - 8 \cdot 04)$ | $(7 \cdot 25 - 7 \cdot 47)$ | $(5 \cdot 58 - 5 \cdot 72)$ | (2.97 - 3.02) | (4.32-4.37) | | |
| Relative risk of mortality | 2.32 | 3.48 | 4.63 | 6.68 | 7.88 | 7.44 | 5.64 | 2.98 | 4.34 | | |
| of sample 2 vs sample 1 | $(2 \cdot 22 - 2 \cdot 42)$ | $(3 \cdot 31 - 3 \cdot 65)$ | $(4 \cdot 45 - 4 \cdot 82)$ | (6-46-6-91) | $(7 \cdot 68 - 8 \cdot 09)$ | $(7 \cdot 29 - 7 \cdot 60)$ | $(5 \cdot 54 - 5 \cdot 74)$ | (2.94 - 3.01) | $(4 \cdot 31 - 4 \cdot 38)$ | | |
| (95% ČI) | | | | | | | | | | | |
| Relative risk of mortality | 2.38 | 3.35 | 4.67 | 6.89 | 7.90 | 7.28 | 5.67 | 3.01 | 4.35 | | |
| of sample 1 vs sample 2 | $(2 \cdot 28 - 2 \cdot 48)$ | $(3 \cdot 19 - 3 \cdot 52)$ | $(4 \cdot 49 - 4 \cdot 86)$ | $(6 \cdot 67 - 7 \cdot 12)$ | $(7 \cdot 70 - 8 \cdot 11)$ | (7.13-7.44) | (5.57-5.77) | (2.98 - 3.047) | $(4 \cdot 32 - 4 \cdot 39)$ | | |
| (95% ĈI) | | | | | | | | | | | |
| Average of two sub- | 2.35 (2.250- | 3.41 (3.250- | 4.65 (4.470- | 6.78 (6.565- | 7.89 (7.690- | 7.36 (7.210- | 5.65 (5.555- | 2.99 (2.960- | 4.34 (4.315- | | |
| sampled relative risks | 2.450) | 3.585) | 4.840) | 7.015) | 8.100) | 7.520) | 5.755) | 3.029) | 4.385) | | |
| *All p-values are <0.001 | | | | | | | | | | | |

| | | 1-fold v | alidation | | 2-fold cros | ss-validation | No validation |
|---------------------|-------------------------------|-----------------------------|----------------------|---------------------------------|---------------------------------------|--|---------------|
| | | | | | Split1 | Split2 | |
| Training set | 80% | 5% | 1% | 0.01% | 50% | 50% | 100% |
| percentage | | | | | | | |
| Training set size | 28,084,615 | 1,753,218 | 349,681 | 3,531 | 17,065,397 | 17,549,071 | 44,913,416 |
| Validation set | 20% | 95% | 99% | 99.99% | 50% | 50% | 0 |
| percentage | | | | | | | |
| Validation set size | 7,014,079 | 33,345,592 | 34,749,098 | 35,095,279 | 17,549,071 | 17,065,397 | 0 |
| Relative risk | 4.34 | 4.28 | 4.43 | 5.58 | 4.37 | 4.33 | 4.35 |
| (Training set) | (4.31-4.37) | $(4 \cdot 16 - 4 \cdot 40)$ | (4.17-4.70) | (3.12-9.98) | (4.33-4.40) | (4.29-4.37) | (4.32-4.37) |
| - | p<0.001 | p<0.001 | p<0.001 | p<0.001 | p<0.001 | p<0.001 | |
| Infection rate | 6.33 | 6.29 | 6.28 | 6.79 | 6.27 | 6.34 | 6.27 |
| (Training set) | (6.32-6.34) | (6.24-6.32) | (6.16-6.32) | (6.22-7.94) | (6.26-6.29) | (6.33-6.35) | (6.264-6.280) |
| Estimated excess | 20,327 | 93,959 | 102,199 | 148,968 | 50,553 | 50,603 | 100,338 |
| COVID-19 death | | | | | | | |
| (Application of RR | | | | | | | |
| and IR from | | | | | | | |
| training set on the | | | | | | | |
| baseline mortality | | | | | | | |
| in validation set) | | | | | | | |
| Observed COVID- | 25,496 | 120,817 | 125,719 | 127,005 | 63,475 | 63,747 | 127,020 |
| 19 death | | | | | | | |
| (Validation set) | | | | | | | |
| Excess/Observed | 0.79 | 0.77 | 0.81 | 1.17 | 0.79 | 0.79 | 0.79 |
| death ratio | | | | | | | |
| Test aim | Test of large training set | Test of underfitting | Test of underfitting | Test of extreme underfitting | Cross-validation of the 50% splitting | Cross-validation of the 50% splitting | No validation |

Table S3 Sensitivity analysis and 2-fold cross-validation of relative risk (RR) and population infection rate (IR)

* **Interpretation:** The training set is used to calculate RR and IR while the validation set is used to estimate baseline 1-year all-cause mortality using KM survival analysis. We applied our model based on RR and IR (from training set) on the baseline mortality (from validation set) to calculate estimated excess COVID-19 deaths. We then compared the estimated excess death to observed excess death in validation set. Due to large number of records in the whole data and randomised splitting without replacement into training and validation sets, even 5% of data as the training set results in close estimations of RR and IR to those of the whole data. To assess any information leak in the analysis pipeline or from training set to validation set, we tested extremely underfitted models. As the size of the training set shrinks below 5%, the training set results in overestimation of RR and IR which demonstrates the lack of information leak from training set to validation are close to the RR of the whole dataset. In our study, we have used the averages of RR and IRs of cross-validated to large data and adequate randomised independently of the splits in this table.

| | | Non-vaco | cinated period | | | Vaccin | ation period | | Mixed period | |
|--|--|--|--|--|------------------------------------|--|------------------------------------|------------------------------------|---------------------------------|--|
| Number of months | 1 st quarter | 2 nd quarter | 3 rd quarter | First 9 months | | 4 th | quarter | | One year | |
| Calendar months | 1 Mar 2020 to 31 May 2020 vs 1 Mar 2018 to 31 May 2018 | 1 Jun 2020 to 31 Aug 2020 vs 1 Jun 2018 to 31 Aug 2018 | 1 Sep 2020 to 30 Nov 2020 vs 1 Sep 2018 to 30 Nov 2018 | 1 Mar 2020 to 30 Nov 2020 vs 1 Mar 2018 to 30 Nov 2018 | | 1 Dec 2020 to 1 Mar 2021 vs 1 Dec 2018 to 1 Mar 2019 | | | | |
| Vaccine dose | - | - | - | - | Overall | 0 dose only | 1 dose only | 2 doses only | Overall | |
| Number of COVID-19* | 184567 | 47264 | 618895 | 850726 | 1290246 | 843451 | 96786 | 2413 | 2140972 | |
| Number of pre- pandemic deaths | 122518 | 108570 | 114353 | 345441 | 133530 | 133530 | 133530 | 133530 | 478971 | |
| Total deaths in pandemic period | 175945 | 110296 | 127453 | 413694 | 171719 | 134307 | 35751 | 1586 | 585413 | |
| Number of deaths with COVID-19 [*] | 33163 | 2219 | 15669 | 60016 | 54802 | 45186 | 8198 | 151 | 127015 | |
| Death-to-exposed (to COVID) ratio | 17.968 | 4.695 | 2.532 | 7.055 | 4.247 | 5.357 | 8.470 | 6.258 | 5.933 | |
| Relative risk (95% CI), p-value | 51.47 (50.90- 52.06) p<0.0001 | 15·12 (14·52-15·76) p<0·0001 | 7·72 (7·59-7·85) p<0·0001 | 7·17 (7·11-7·23) p<0·0001 | 11.05 (10.95-11.16) p<0.0001 | 13·94 (13·80-14·09) p<0·0001 | 22.05 (21.58-22.52) p<0.0001 | 16·29 (13·96-19·01) p<0·0001 | 4·35 (4·32-4·37) p<0·0001 | |
| Rate [*] ratio (95% CI), p-value | 55.85 (53.19- 58.64) p<0.0001 | 15.52 (13.83-17.42) p<0.0001 | 7·76 (7·50-8·03) p<0·0001 | 7·42 (7·29-7·55) p<0·0001 | 11·24 (10·98-11·52) p<0·0001 | 14·28 (13·88-14·69) p<0·0001 | 22.52 (21.01-24.15) p<0.0001 | 16·56 (10·58-25·91) p<0·0001 | 4·42 (4·38-4·46) p<0·0001 | |
| Infection risk % (95% CI) | 0·54 (0·538- 0·543) | 0·14 (0·138-0·140) | 1.83 (1.824-1.833) | 2·49 (2·487-2·498) | 3.83 (3.820-3.833) | 4.64 (4.626-4.646) | 2·94 (2·927-2·944) | 1·4 (1·367-1·427) | 6·27 (6·264-6·281) | |

Table S4 Assessment of risk ratio, rate ratio, infection risk, and infection rate of different periods based on vaccination in people of age 30 or older

*When the vaccine dose is taken into account, only COVID-19 and deaths after vaccination with the specified dose is included in the analysis

**The denominator for rate is expressed as person-years-

Interpretation: The low IR values for pre-vaccination period, especially 1Mar-31Aug, which is also evident in https://coronavirus-data-gov-uk/details/cases, could be attributed to data collection and testing methods in the UK. These low numbers affect the overall IR results, causing an underestimation of IR. The values of relative risk or rate ratio in the vaccination period should be interpreted in the context of the study design. In our study, the COVID-unexposed group is from Mar2018-Mar2019 where COVID-19 vaccination was not meaningful; therefore, while the denominator of risk in exposed group decreases (due to narrowing the cohort to specific doses), the denominator of the risk in unexposed group does not change, resulting in large numeric results. One potential approach to address this issue is one-to-one matching between vaccinated exposed people and unexposed people from the pre-pandemic period, which is beyond the scope of our study. However, unlike the quarterly analysis, the overall 1-year RR is based on the total denominator. The ratio of death to the number of exposed to COVID is a better measure to compare different doses of vaccination with regards to COVID-19 infection and mortality. We do not have any information on the actual onset of COVID-19 symptoms, infection rate, or immunity level before COVID-19 infection in vaccinated people. The comparison of infection rate between 0, 1 and 2 doses show that people with 1 or 2 doses of vaccinate have lower infection rates comparing with people without any vaccine. Although, the inclusion of people with 1 or 2 doses causes a slight decrease in the overall IR (from 4-64 to 3-83) in the 4th quarter, the difference is negligible compared to the effect of low IR in the first stages of the pandemic.

| Underlying | Age \leq 70 years | | | Age > 70 years | | | All ages | | |
|---|---------------------|-----------------|---------------------------------------|-----------------|-----------------|-------------------------------------|---------------------|-----------------|-------------------------------------|
| conditions | N (%) | Observed deaths | 1-year mortality risk % (95% CI) * | N (%) | Observed deaths | 1-year mortality risk % (95% CI) | N (%) | Observed deaths | 1-year mortality risk % (95% CI) |
| At least one comorbidity except for age > 70 | 4744687 (13.52) | 54497 | 1.15 (1.14-1.16) | 3845654 (10.96) | 250715 | 6.52 (6.49-6.54) | 8590341 (24·47) | 305121 | 3.55 (3.54-3.57) |
| Age > 70 | - | - | - | 7048826 (20.08) | 374134 | 5.31 (5.29-5.32) | - | - | - |
| Diabetes | 1681745 (4.79) | 17363 | 1.03 (1.02-1.05) | 1247763 (3.55) | 79630 | 6.38 (6.34-6.42) | 2929508 (8.35) | 96993 | 3.31 (3.29-3.33) |
| CVD | 1382451 (8.93) | 23139 | 1.67 (1.65-1.70) | 2049396 (5.84) | 171361 | 8.36 (8.32-8.40) | 3431847 (9.78) | 194500 | 5.67 (5.64-5.69) |
| BMI > 40 | 934564 (2.66) | 24 | 0.003 (0.002-0.004) | 286124 (0.82) | 62 | 0.02 (0.02-0.03) | 1220688 (3.48) | 86 | 0.007 (0.006-0.009) |
| Steroid therapy | 1945757 (5.54) | 22317 | 1.15 (1.03-1.16) | 1069149 (3.05) | 57443 | 5.37 (5.33-5.42) | 3014906 (8.59) | 79760 | 2.65 (2.63-2.66) |
| COPD | 578581 (1.65) | 14227 | 2.46 (2.42-2.50) | 701147 (2.00) | 64778 | 9.24 (9.17-9.31) | 1279728 (6.65) | 79005 | 6.17 (6.13-6.22) |
| CKD | 502776 (1.43) | 4201 | 0.84 (0.81-0.86) | 1204554 (3.43) | 40589 | 3.37 (3.34-3.40) | 1707330 (4.86) | 44790 | 2.62 (2.60-2.65) |
| Chronic liver disease | 68596 (0.19) | 4405 | 6.42 (6.24-6.60) | 20841 (0.06) | 2854 | 13.69 (13.23-14.16) | 89437 (0.25) | 57259 | 8.12 (7.94-8.30) |
| Number of underlying conditions | | | | | | | | | |
| 3+ | 150113 (0.71) | 7418 | 2.97 (2.90-3.03) | 530160 (1.51) | 39374 | 7.43 (7.36-7.50) | 780273 (2.22) | 46792 | 6.00 (5.94-6.05) |
| 2 | 857996 (2.44) | 14535 | 1.69 (1.67-1.72) | 1114997 (3.18) | 77778 | 6.98 (6.93-7.02) | 1972993 (5.62) | 92313 | 4.68 (4.65-4.71) |
| 1 | 3636578 (10.36) | 32544 | 0.89 (0.89-0.90) | 2200497 (6.27) | 133563 | 6.07 (6.04-6.10) | 5837075 (16·63) | 166107 | 2.85 (2.83-2.86) |
| 0 | 23305297 (66.40) | 50340 | 0.22 (0.21-0.22) | 3203172 (9.13) | 123419 | 3.85 (3.83-3.87) | 26508469 (75·52) | 173759 | 0.655 (0.652-0.659) |

Table S5 Baseline one year mortality risk in England per underlying condition (NHS Digital TRE; n = 35,098,810 age ≥30 years)

* Risk of death per 100 based on Kaplan-Meier estimate of one year mortality TRE: Trusted Research Environment

| | No CVD | CVD | No DM | DM | No CKD | CKD | No COPD | COPD |
|---------------|---------|--------|---------|--------|---------|--------|---------|--------|
| CVD | - | - | 2645479 | 786368 | 2802686 | 629161 | 2979024 | 452823 |
| Diabetes | 2143140 | 786368 | - | - | 2446463 | 483045 | 2665508 | 264000 |
| CKD | 1078169 | 629161 | 1224285 | 483045 | - | - | 1515724 | 191606 |
| COPD | 826905 | 452823 | 1015728 | 264000 | 1088122 | 191606 | - | - |
| BMI>40 | 1030358 | 190330 | 909196 | 311492 | 1086648 | 134040 | 1141735 | 78953 |
| Chronic liver | 64493 | 24944 | 63228 | 26209 | 80570 | 8867 | 75949 | 13488 |
| disease | | | | | | | | |
| Steroid | 2429993 | 584913 | 2607609 | 407297 | 2710601 | 304305 | 2372332 | 642574 |
| therapy | | | | | | | | |

Table S6 Overlap between high-risk groups for CVD, DM, CKD, and COPD in England (NHS Digital TRE; n = 35,098,810 age ≥30)

*CVD: cardiovascular disease; DM: diabetes mellitus; CKD: chronic kidney disease; COPD: chronic obstructive pulmonary disease