

开放
数据

我国省级地方政府新冠疫情 数据发布研究报告



报告制作方 | 复旦大学数字与移动治理实验室
DATAMATE数据工作室



 背景与目的

 数据采集与分析

 研究发现

 讨论与调查

 结论与建议

 附录

背景与目的

新冠疫情暴发以来，我国各地政府部门通过多种渠道发布疫情信息，包括领导重要指示、政策文件通知、防控工作动态、自我防护知识等内容，这些信息有助于公众获知疫情的最新情况和相关知识，在一定程度上满足了公众的信息需求。

《中华人民共和国传染病防治法》第三十八条规定“公布传染病疫情信息应当及时、准确。”2020年2月23日，习近平总书记在统筹推进新冠肺炎疫情防控和经济社会发展工作部署会议上要求，“要完善疫情信息发布，依法做到公开、透明、及时、准确”。

进入大数据时代，公众的信息需求也已发生了变化，除了以上这些疫情信息和知识，公众还想获得一些更为具体的、详细的、动态的数据。比如，有关一个地方疫情概况的最新数据，今天有多少疑似和确诊病例？新增了多少？累计有多少？这些病人目前情况如何？有多少危重？有多少死亡或出院？又比如，有关病人个人的具体数据，多大年龄？去过哪里？从哪里来？坐什么交通工具来？住在哪里？接触过什么人？在哪家医院就诊的？

根据《中华人民共和国传染病防治法》第三十八条“传染病暴发、流行时，国务院卫生行政部门负责向社会公布传染病疫情信息，并可以授权省、自治区、直辖市人民政府卫生行政部门向社会公布本行政区域的传染病疫情信息。”

因此，报告对我国31个省、自治区、直辖市（不含港澳台地区，以下简称各省份）人民政府的卫生健康委员会在新冠疫情暴发后发布的疫情数据进行了系统的采集、分析和比较，包括其发布内容、发布形式和发布时间等方面，并总结规律，发现问题和提出建议。

数据采集与分析

本报告主要采集了以下数据并进行分析。首先，报告采集了我国各省级地方政府卫生健康委员会官方网站上发布的疫情通报类信息，例如在标题中出现“疫情通报”“疫情情况”“病例”“个案信息”等字样，并在内容中发布地方疫情的统计数据 and 病例个人情况的信息。数据采集时段为从疫情开始到 2020 年 2 月 10 日 24 时，这一时段基本覆盖了各省份从疫情发生到疫情高峰的时期，是社会公众对疫情数据的需求最大最迫切的时期，也是检验一个地方疫情数据发布水平最关键的时期。

报告对各地方政府疫情数据发布的整体分析框架见表 1，包括整体发布情况、疫情统计数据发布情况、病例个案信息发布情况等方面。然后，报告对各省份疫情通报中发布的疫情统计数据（截至 2020 年 2 月 10 日 24 时，共 891 条）和病例个案数据（截至 2020 年 2 月 5 日 24 时，共 921 例）进行了编码与统计分析。

表 1 整体分析框架

一级维度	二级维度
整体发布情况	时效性
	易得性
	可读性
疫情统计数据发布情况	全面性
	持续性
	覆盖层级
病例个案信息发布情况	公开率
	精细度
	隐私保护

其次，报告对当前互联网上对疫情数据发布的观点进行了采集与分析。以“病例信息公开”“病例轨迹公开”“病例行踪公开”等作为关键词句进行检索，在人民日报、新华社等公众媒体，微博、微信等社交媒体网络平台上，就“病例信息应该开放到多细”这一话题收集相关报道与社会讨论，数据采集时间截止到2020年2月15日。报告对采集到的数据进行了文本分析和梳理提炼，并选取了具有代表性的专家评论、媒体观点、网友点评进行例举。

再次，针对“疫情数据应该发布到多细”这一问题，报告于2020年2月16日起通过“复旦DMG”微信公众号进行了初步的网络调查（问卷见附录）。截至2020年2月29日，共收到反馈问卷576份，报告对调查结果进行了统计分析。

最后，基于以上各项分析，报告对我国政府的疫情数据发布提出政策建议。





研究发现

1. 整体发布情况

发布疫情通报应及时、动态、便于公众获取和解读。因此，报告首先从时效性、易得性、可读性三个方面对各省级政府疫情数据发布的整体情况进行了分析。

1.1 时效性

报告对各省份卫健委发布疫情数据的时效性进行了研究，包括发布起始日期和平均每日发布次数两个方面。各省份开始发布疫情数据的日期如图 1 所示。上海、广东开始日期最早（1 月 20 日）；西藏开始日期最晚（1 月 30 日）；其余省份开始发布疫情数据的日期都在 1 月 21 日至 1 月 24 日之间，其中，1 月 21 日有 11 个省份开始发布，包括湖北省。除湖北之外，绝大多数省份发布疫情数据的开始日期都基本与出现首个确诊病例的日期同步。

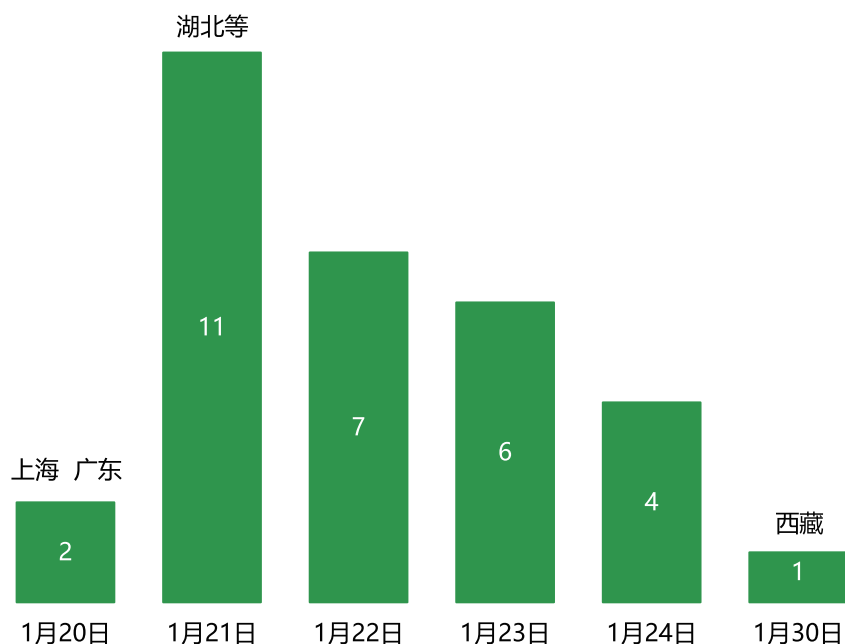


图 1 疫情数据发布起始日期与省份数量

各省份疫情数据平均每日发布的次数如图 2 所示。重庆平均每日发布次数最多（2.9 次），在报告观测期内，每日根据疫情发展多次发布通报，灵活机动；宁夏的平均每日发布次数最少（低于 1 次）。在报告观测期内，多数省份的平均每日发布次数分布在 1 次左右。

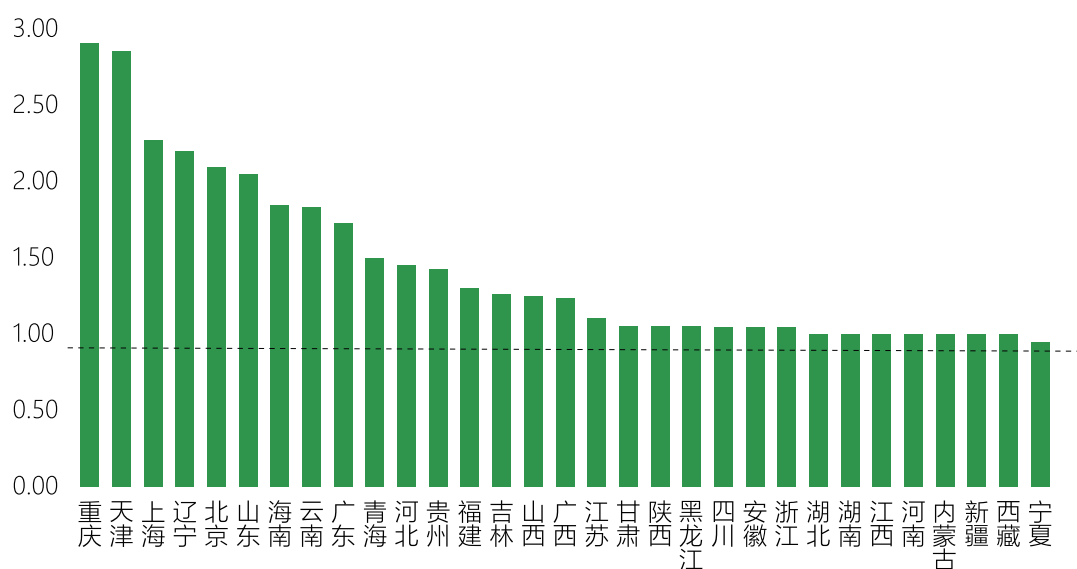


图 2 各省份疫情数据平均每日发布次数

1.2 易得性

报告对获得各省份疫情数据的便利性进行了研究，主要观察各省份卫健委在官网首页设置疫情防控专题、在专题下设置疫情通报专栏、以及疫情通报的标题形式等方面。

1.2.1 疫情防控专题设置

在报告观测期内，90% 的省份都在首页开通了与疫情防控相关的专题，对疫情发展情况、疫情防控工作动态、新闻报道、防控知识等内容进行专门发布，便于公众访问相关内容，如以下截图。其余的省份尚未在首页上开通相关专题，疫情通报数据通常与其他新闻混杂在一起发布，不便于公众获取。



图 3 北京市卫健委官网首页截图
(截图于 2020.1.28)

各省份在卫健委官网首页设置疫情防控专题的情况如图 4 所示，除江苏、黑龙江和西藏外，其他省份都设置了专题。

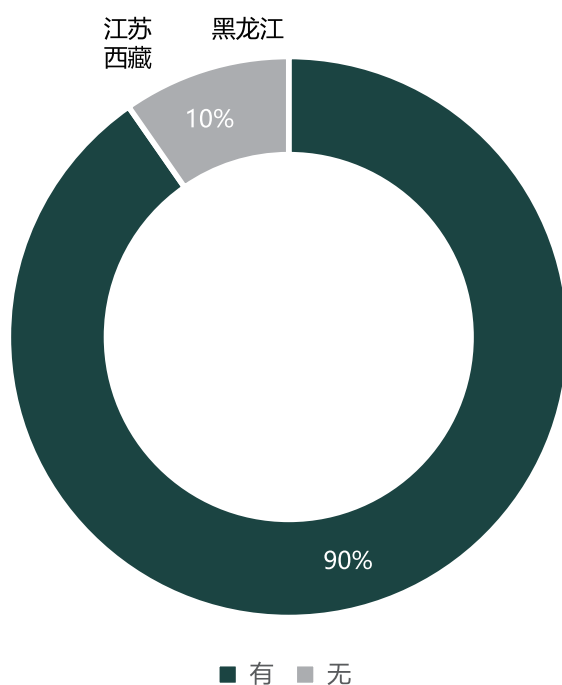


图 4 各省份在卫健委官网首页设置疫情防控专题情况

1.2.2 疫情通报专栏设置

在报告观测期内，61% 的省份还在疫情防控相关专题下开通了与疫情通报相关的专栏，集中发布疫情数据，便于公众获取（如以下截图）。其余省份则多将疫情数据与防控工作动态、新闻通告等信息混杂在一起进行发布，不便于公众访问和获取。



图 5 天津市卫健委官网“疫情通报”专栏截图
(截图于 2020.1.28)

各省份在卫健委官网首页疫情防控专题下设置疫情通报专栏的情况如图 6 所示。

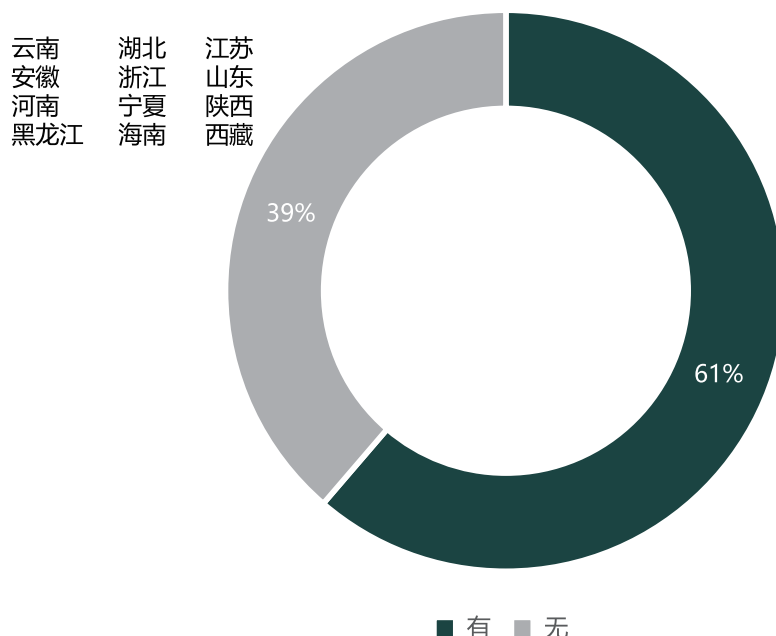


图 6 各省份在疫情防控专题下设置疫情通报专栏情况

1.2.3 疫情通报标题形式

发布疫情通报时使用的标题名称也会影响公众访问和获取相关数据的便利性。报告发现目前各省份发布疫情通报时主要使用两种形式的标题名称：一类是统一标准类，即每日统一以标准化的标题进行发布，例如重庆市卫健委发布的疫情通报标题：“2020 年 1 月 29 日重庆市新型冠状病毒感染的肺炎疫情情况”。另一类是实时数据类，以在标题中直接呈现“确诊病例”“累计病例”“死亡病例”“出院病例”实时数据的方式来通报疫情，例如北京市卫健委发布的疫情通报标题：“1 月 28 日 12 时至 29 日 12 时，本市新增 11 例新型冠状病毒感染的肺炎病例”。

总体上，统一标准类标题规范性强，便于用户查找，但未能在标题中实时展现关键数据，动态性相对较弱；而实时数据类标题能在标题中直接呈现关键数据，时效性强，灵活机动，但缺乏统一规范，不便于用户搜索查找。

在报告观测期内，各省份疫情通报标题形式的使用情况如图 7 所示。16% 的省份只采用统一标准类标题；13% 的省份只采用实时数据类标题；其余 71% 的省份则两类标题形式混合使用。

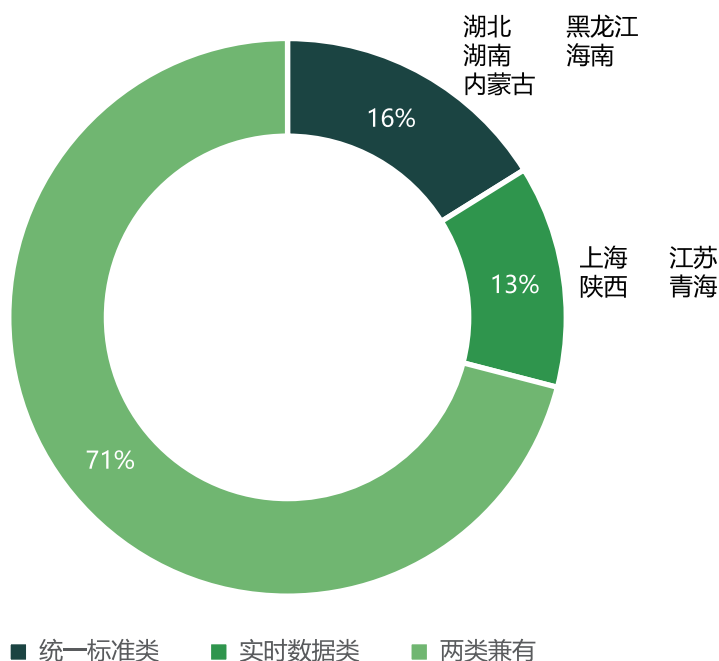


图 7 各省份疫情通报标题形式使用情况

1.3 可读性

报告对各省份疫情数据的发布形式进行了分析。在报告观测期内，各省份主要采用了文字、表格和可视化图形等形式，给公众带来的阅读体验和解读难度也各不相同。

目前，所有省份发布的疫情通报都存在“以文字夹杂数字”的形式，如以下截图。

2020年1月29日重庆市新型冠状病毒感染的肺炎疫情情况

发布时间：2020-01-29	来源：宣传统战部	浏览量：4896次	编辑：办公室	审核：宣传统战部	字体大小： 小 中 大	
2020年1月28日0—24时，重庆市报告新型冠状病毒感染的肺炎新增确诊病例15例，新增重型病例3例，重型病例转普通型病例1例，新增危重型病例1例。渝中区为报告首例确诊病例。 新增确诊病例中，万州区2例、渝中区1例、江北区1例、合川区1例、綦江区2例、璧山区1例、开州区2例、忠县2例、云阳县1例、奉节县1例、石柱县1例；新增重型病例中，渝北区1例、江津区1例、璧山区1例；重型病例转普通型病例1例为永川区；新增危重型病例1例为铜梁区。 截至1月28日24时，重庆市累计报告新型冠状病毒感染的肺炎确诊病例147例，重型病例15例，危重型病例4例。 确诊病例中，万州区17例、黔江区1例、涪陵区1例、渝中区1例、大渡口区6例、江北区5例、九龙坡区5例、南岸区1例、渝北区6例、巴南区1例、长寿区7例、江津区2例、合川区5例、永川区3例、綦江区4例、大足区1例、璧山区4例、铜梁区2例、潼南区2例、开州区11例、梁平区1例、城口县1例、丰都县3例、垫江县5例、忠县11例、云阳县11例、奉节县4例、巫山县5例、巫溪县5例、石柱县7例、秀山县1例、两江新区8例；重型病例中，万州区2例、大渡口区2例、江北区1例、九龙坡区1例、渝北区1例、长寿区2例、江津区1例、璧山区2例、垫江县1例、忠县1例、云阳县1例；危重型病例中，万州区1例、江北区1例、九龙坡区1例、铜梁区1例。 累计追踪到密切接触者2462人，已解除医学观察103人，尚有2359人正在接受医学观察。						

图 8 重庆市卫健委官网疫情情况截图
(截图于 2020.1.29)

以这种形式发布疫情数据，还不够清晰直观，缺乏数据思维，不便于用户理解和分析，如用户想对这些数据进行进一步分析，需要先将数据从这些文字中提取和整理出来，做成数据表格。

因此，为了方便公众解读和分析数据，19% 的省份在文字形式之外，开始增加了结构化的表格形式来呈现疫情数据，如以下截图。

2020年2月8日12—24时，重庆市报告新型冠状病毒肺炎（简称新冠肺炎）新增确诊病例18例，新增重型病例4例，重型病例转普通型病例4例，重型病例转危重型病例3例，危重型病例转重型病例1例，新增危重型病例1例，新增出院病例8例。

新增确诊病例中，万州区6例、江北区3例、长寿区3例、合川区1例、永川区2例、铜梁区1例、垫江县1例、巫溪县1例；新增重型病例中，万州区1例、合川区1例、潼南区1例、奉节县1例；重型病例转普通型病例中，璧山区1例、开州区1例、忠县1例、奉节县1例；重型病例转危重型病例中，江北区1例、九龙坡区1例、渝北区1例；危重型病例转重型病例为忠县1例；新增危重型病例为万州区1例；新增出院病例中，万州区1例、渝中区1例、九龙坡区1例、梁平区1例、忠县2例、云阳县2例。

截至2月8日24时，重庆市累计报告新冠肺炎确诊病例446例；现有重型病例29例，危重型病例20例，死亡病例2例，出院病例39例。累计追踪到密切接触者14765人，已解除医学观察8073人，尚有6692人正在接受医学观察。具体情况如下：

序号	区域	确诊病例		现有重型病例	现有危重型病例	死亡病例	出院病例
		新增	累计				
	全市	18	446	29	20	2	39
1	万州区	6	84	12	5	1	5
2	黔江区		2				
3	涪陵区		3				
4	渝中区		20				1
5	大渡口区		7				1
6	江北区	3	23	2	3		1
7	沙坪坝区		2				
8	九龙坡区		19		2	1	1
9	南岸区		8		1		
10	北碚区						
11	渝北区		12		1		2
12	巴南区		3				
13	长寿区	3	15	1			2
14	江津区		4				1
15	合川区	1	17	1			1
16	永川区	2	5				1
17	南川区						
18	綦江区		15	1	1		
19	大足区		10				
20	璧山区		8				
21	铜梁区	1	8	2			
22	潼南区		13	1			
23	荣昌区		8				
24	开州区		18	1	1		2
25	梁平区		4				1
26	武隆区		1		1		
27	城口县		2				1
28	丰都县		9	1	1		1
29	垫江县	1	19				1
30	忠县		16	3			3
31	云阳县		20		1		7
32	奉节县		14	3	2		1
33	巫山县		8				4
34	巫溪县	1	12				1
35	石柱县		12				
36	秀山县		1				1
37	酉阳县		1				
38	彭水县		2				
39	两江新区		16	1	1		
40	高新区		4				
41	万盛经开区		1				

图 9 重庆市卫健委发布的疫情数据截图
(截图于 2020.2.9)

还有 23% 的省份采用了可视化图形的形式，通过统计图表、数据地图等方式来直观清晰地展现疫情分布和发展趋势等信息，使公众一目了然（如图 10）。

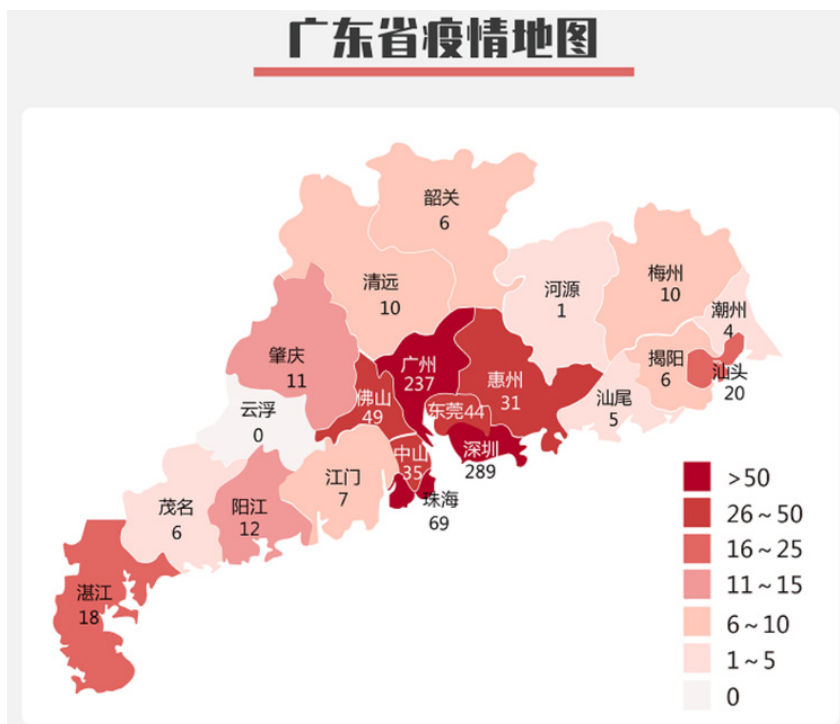


图 10 广东省卫健委发布的疫情数据截图
(截图于 2020.2.5)

在报告观测期内，各省份采用文字、表格和可视化图形发布疫情数据的情况如图 11 所示。



图 11 各省份文字、表格、可视化图形采用情况

2. 疫情统计数据发布情况

面对疫情，公众希望获知有关一个地方疫情概况的统计数据：比如各个省市今天累计有多少疑似、确诊、危重、出院或死亡病例？今天又新增或减少了多少？面对以上需求，报告发现各省份早期发布的主要是本地的疫情统计数据，包括疑似病例、确诊病例、密切接触者、病情平稳者病例、重症危重者病例、死亡病例、出院病例等数据（如下图）。

上海新增11例新型冠状病毒感染的肺炎确诊病例

(2020-02-03)

2020年2月2日12—24时，上海市排除新型冠状病毒感染的肺炎疑似病例60例；新增确诊病例11例，其中外地来沪人员2例，本市常住人口9例。具体情况如下：

区域		确诊病例
外地来沪人员 (按居住地分)	湖北武汉	2
	浦东	4
本市常住人口 (按居住地分)	杨浦	2
	宝山	2
	普陀	1
	合计	11

截至2月2日24时，上海市已累计排除疑似病例449例，发现确诊病例193例。确诊病例中，男性101例，女性92例；年龄最大88岁，最小7岁；113例有湖北居住或旅行史，8例有湖北以外地区居住或旅行史，70例有相关确诊病例接触史，2例无湖北接触史；外地来沪人员76例，本市常住人口117例。具体情况如下：

区域		确诊病例
外地来沪人员 (按居住地分)	湖北武汉	54
	湖北其他地市	20
	安徽	1
	黑龙江	1
本市常住人口 (按居住地分)	浦东	39
	宝山	10
	静安	9
	闵行	9
	徐汇	7
	长宁	7
	松江	7
	黄浦	5
	虹口	5
	奉贤	5
	普陀	4
	杨浦	4
	嘉定	2
	青浦	2
	金山	1
	崇明	1
	合计	193

目前，170例病情平稳，7例病情危重，5例重症，10例治愈出院，1例死亡。尚有173例疑似病例正在排查中。

图 12 上海市卫健委官网疫情情况截图
(截图于 2020.2.3)

2.1 全面性

目前，各省份公布的各类疫情统计数据基本上覆盖了疫情发展的全周期，包括了一个病患从疑似、确诊（或被排除）、经治疗后病情平稳到出院（或是转为重症、危重或不幸死亡）的各个环节；同时还包括了对确诊病人的密切接触者进行医疗观察，以及确诊（或解除观察）的过程，各个环节之间的顺序和关系如图 13 所示。

报告基于图 13 中的各个环节，对各省份卫健委官网上发布的疫情统计数据的全面性进行了分析。分析发现，各省份发布疫情统计数据的全面程度并不一致，有些数据在大多数省份都公布过，而有些数据只在个别省份公布过。图 13 中标注了各项数据在所有省份的整体公布比例（只要某个省份在 1 月 20 日到 2 月 10 日报告观测期内公布过一次该项数据，即视为曾经公布过）。



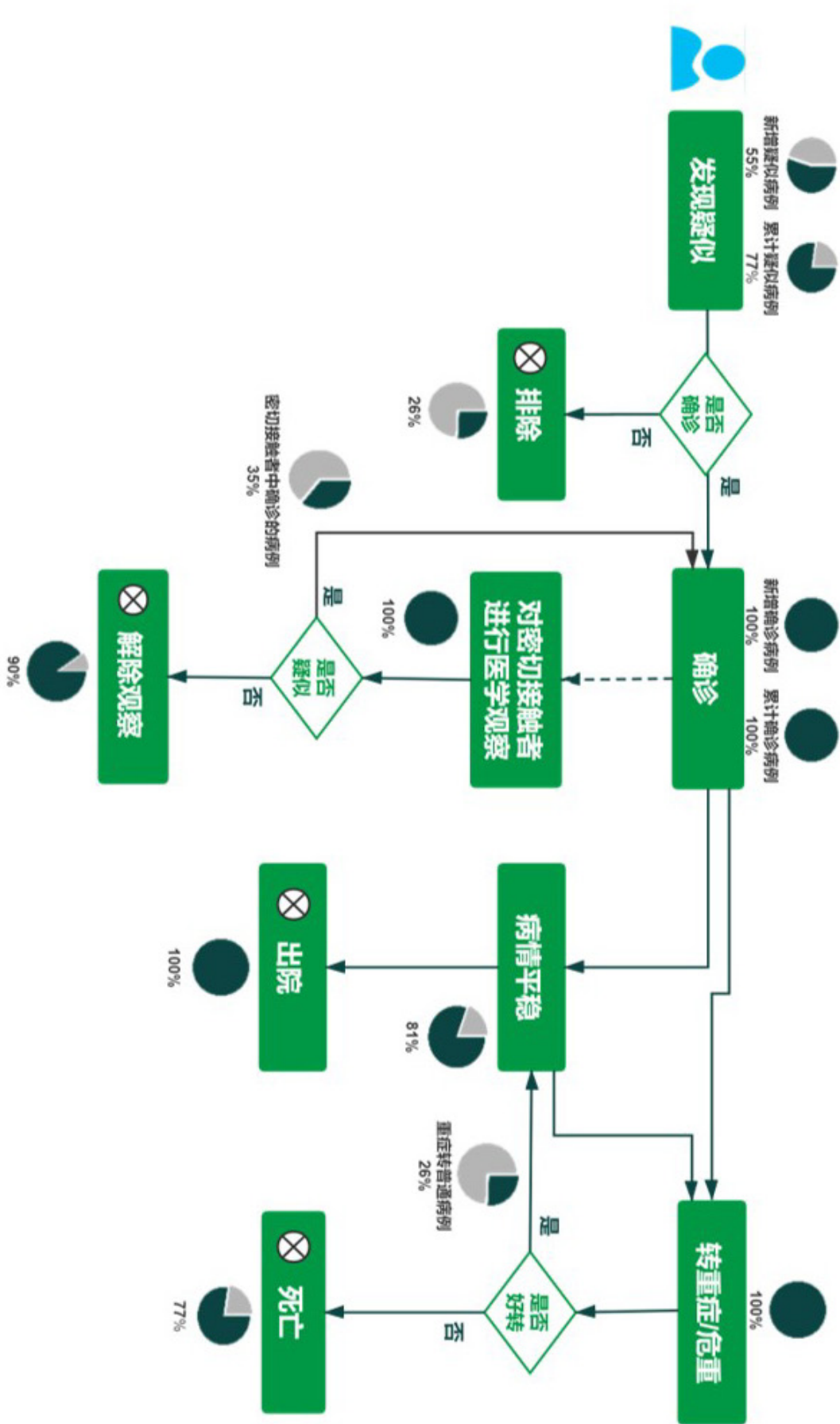


图 13 疫情统计数据发布的全面性

从图中可见，在报告观测期内，新增确诊病例、累计确诊病例、重症 / 危重病例、接受医学观察者数量、出院病例这五项数据所有省份都发布过，比例达到 100%；其次是解除医学观察数量、病情平稳病例、死亡病例、累计疑似病例、新增疑似病例等数据，都有超过 50% 的省份发布过；而排除疑似病例数、重症转为普通型病例这两项数据只在不到 30% 的省份发布过。

图 14 具体展示了各省份发布的疫情统计数据全面性（在报告观测期内，只要该省份公布过一次该项数据即被视作曾经公布过）。从图中可见，在报告观测期内，重庆、吉林与山西发布疫情统计数据的全面性最高，覆盖了所有 14 项主要数据；其次是广西，发布了 13 项数据；而江西与新疆只发布了 7 项数据。



图 14 各省份疫情统计数据发布的全面性

2.2 持续性

各项疫情统计数据应当在疫情期间持续发布。然而，在报告观测期内，没有一个省份连续发布了每项统计数据，有些省份曾发布过某项数据，但之后不再发布，或只是断断续续发布；有些省份则从未发布过某项数据。各省份发布某项数据的天数各不相同，还没有统一的标准和规范。

图 15 是 1 月 20 日到 2 月 10 日报告观测期内，各项疫情统计数据每日在所有省份的发布比例，颜色越深代表当天有更高比例的省份发布了相应的数据。例如，1 月 30 日有 97% 的省份发布了“累计确诊病例”这项数据。

从图中可以发现，在报告观测期内，随着疫情的发展，全国各省份发布的疫情统计数据不断丰富全面，其中累计确诊病例、新增确诊病例、正在接受医学观察者、累计重症 / 危重病例等数据每日在全国各省份发布的比例最高，而排除疑似病例与重症转普通型病例两项数据的比例最低。

	累计确诊病例	新增确诊病例	正在接受医学观察者	累计重症/危重病例	新增或累计出院病例	已解除医学观察者	新增或累计死亡病例	累计疑似病例	新增/累计疑似转确诊病例	新增疑似病例	密切接触者中确诊病例	新增或累计病例中有湖北武汉接触史	排除疑似病例	累计重症转为普通型病例	平均比例
02-10	97%	90%	77%	84%	90%	71%	68%	52%	29%	39%	19%	10%	16%	0%	53%
02-09	97%	90%	81%	90%	94%	74%	71%	55%	26%	35%	19%	10%	13%	6%	54%
02-08	97%	94%	84%	90%	94%	77%	68%	58%	32%	39%	23%	10%	10%	3%	56%
02-07	100%	94%	84%	94%	94%	74%	71%	48%	29%	32%	16%	10%	13%	3%	54%
02-06	97%	87%	84%	90%	90%	71%	68%	45%	29%	32%	16%	10%	13%	6%	53%
02-05	94%	87%	81%	87%	90%	65%	65%	45%	29%	29%	23%	10%	10%	6%	51%
02-04	97%	90%	84%	94%	94%	68%	61%	48%	32%	32%	23%	10%	6%	3%	53%
02-03	100%	94%	87%	94%	90%	71%	65%	52%	29%	32%	23%	10%	13%	6%	55%
02-02	100%	97%	87%	100%	87%	65%	68%	42%	35%	39%	23%	13%	6%	3%	55%
02-01	97%	97%	84%	90%	81%	58%	68%	42%	32%	29%	26%	16%	10%	6%	53%
01-31	97%	94%	81%	94%	81%	58%	65%	42%	29%	29%	23%	13%	6%	3%	51%
01-30	97%	100%	77%	87%	71%	61%	52%	35%	29%	23%	19%	16%	3%	6%	48%
01-29	97%	97%	84%	81%	65%	65%	52%	32%	23%	23%	23%	13%	6%	6%	47%
01-28	97%	97%	90%	74%	42%	61%	42%	29%	23%	16%	23%	16%	3%	6%	44%
01-27	97%	97%	87%	77%	32%	55%	48%	29%	23%	16%	19%	16%	6%	3%	43%
01-26	97%	97%	84%	65%	26%	52%	35%	26%	16%	13%	10%	13%	0%	6%	38%
01-25	90%	94%	81%	58%	32%	45%	39%	23%	26%	16%	10%	19%	0%	3%	38%
01-24	84%	84%	77%	48%	26%	35%	29%	16%	23%	13%	0%	16%	3%	3%	33%
01-23	55%	68%	48%	29%	10%	13%	16%	13%	32%	6%	3%	32%	3%	0%	24%
01-22	45%	52%	29%	19%	3%	3%	10%	13%	42%	13%	3%	29%	3%	0%	19%
01-21	26%	39%	26%	10%	3%	3%	6%	19%	29%	3%	3%	19%	0%	0%	13%
01-20	0%	10%	6%	0%	0%	0%	0%	0%	3%	0%	0%	3%	0%	0%	2%
平均比例	84%	84%	73%	71%	59%	52%	48%	35%	27%	23%	16%	14%	7%	4%	43%

图 15 疫情统计数据在所有省份的每日发布比例

报告还进一步分析了各省份发布各项数据的持续性，以如下方式计算：

某地发布该项数据的持续性 = 累计发布该项数据的天数 / 观测总天数

图 16 展示了各省份发布各项统计数据的持续性，图中某个方框的颜色越深，比例数值越高，则表明该省份持续发布该项数据的天数越多，中断天数越少。比较发现，在报告观测期内，上海和山东发布各项数据的整体持续性最高，分别为 62% 和 58%；其次是福建、贵州、黑龙江等地，都在 55% 左右；相对而言，西藏、宁夏、内蒙古等地的持续性最低。

数据发布的持续性也反映了一个省份数据发布的规范性。例如，上海在当天并无重症病例出现的情况下，仍然坚持发布该项数据，只是将该数据的数值标为“无”，体现了数据发布的规范性和持续性。而有些省份未能每天持续发布该项数据，使公众无法区分当天到底是数据缺失，即没有发布该项数据，还是当天该项数据的数值为 0。



	新增疑似病例			累计疑似病例			排除疑似病例			新增确诊病例			累计确诊病例			新增或累计病例中有湖北/武汉接触史			正在接受医学观察者			已解除医学观察者			密切接触者中确诊的病例			新增/累计病情平稳病例			累计重症/危重病例			新增或累计死亡病例			新增或累计出院病例			平均比例																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
上海	0%	95%	68%	100%	95%	82%	9%	0%	73%	95%	95%	0%	73%	82%	62%																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		</

图 16 各省份发布疫情统计数据的持续性比例

2.3 覆盖层级

报告发现，有些省份的卫健委平台只发布了省一级的数据，而有些省份则同时发布了下辖各市（区）的数据。报告以公众最为关注的新增确诊病例和累计确诊病例数据为例，分析了1月20日到2月10日报告观测期内各省份发布这两项数据时覆盖到省级和市级（或区级）的比例。

分析结果如图17所示，图中某个方框的颜色越深，则表示当天有更高比例的省份发布了该项数据。例如，1月21日有16%的省份发布了“新增确诊病例（各市/区）”这项数据。总体上，报告观测期内，随着时间的推移，各省份发布的疫情统计数据覆盖其下辖各市（区）数据的比例越来越高，但整体上发布下辖各市（区）数据的比例低于发布省级数据的比例。



	累计确诊病例 (全省)	新增确诊病例 (全省)	累计确诊病例 (各市/区)	新增确诊病例 (各市/区)	平均比例
02-10	97%	90%	90%	77%	89%
02-09	97%	90%	94%	77%	90%
02-08	97%	94%	94%	77%	90%
02-07	100%	94%	97%	81%	93%
02-06	97%	87%	90%	81%	89%
02-05	94%	87%	90%	77%	87%
02-04	97%	90%	94%	81%	90%
02-03	100%	94%	97%	81%	93%
02-02	100%	97%	97%	77%	93%
02-01	100%	97%	97%	77%	93%
01-31	97%	94%	94%	71%	89%
01-30	97%	100%	97%	71%	91%
01-29	97%	97%	97%	74%	91%
01-28	97%	97%	97%	77%	92%
01-27	97%	97%	97%	81%	93%
01-26	97%	97%	90%	61%	86%
01-25	90%	94%	81%	55%	80%
01-24	84%	84%	68%	55%	73%
01-23	52%	65%	35%	42%	48%
01-22	45%	52%	29%	19%	36%
01-21	26%	39%	13%	16%	23%
01-20	0%	10%	0%	3%	3%
平均比例	84%	84%	79%	64%	78%

图 17 省市两级数据在所有省份的每日发布比例

图 18 展示了各省份发布省级和市级（或区级）新增和累计确诊病例数据的持续性，方框颜色越深，比例数值越高，则表示该省份持续发布该项数据的天数越多，中断天数越少。从图中可见，报告观测期内，海南与山东发布省市两级数据的比例最高（100%），其次为江苏、安徽、浙江和重庆，都 98% 以上。



	新增确诊病例 (全省)	新增确诊病例 (各市/区)	累计确诊病例 (全省)	累计确诊病例 (各市/区)	平均比例
海南	100%	100%	100%	100%	100%
山东	100%	100%	100%	100%	100%
江苏	100%	100%	100%	94%	99%
安徽	95%	100%	95%	100%	98%
浙江	95%	95%	100%	100%	98%
重庆	100%	90%	100%	100%	98%
江西	100%	95%	100%	95%	97%
新疆	100%	95%	100%	95%	97%
广东	100%	100%	95%	91%	97%
湖北	100%	100%	100%	86%	96%
河南	100%	90%	95%	95%	95%
山西	100%	89%	95%	95%	95%
福建	100%	95%	90%	90%	94%
云南	100%	72%	100%	100%	93%
甘肃	100%	86%	93%	93%	93%
湖南	100%	100%	86%	86%	93%
河北	100%	89%	89%	89%	92%
广西	100%	71%	100%	95%	92%
吉林	100%	74%	95%	95%	91%
陕西	100%	79%	95%	79%	88%
辽宁	80%	80%	100%	90%	88%
天津	100%	70%	100%	75%	86%
贵州	95%	62%	90%	86%	83%
青海	94%	50%	89%	83%	79%
四川	90%	33%	95%	95%	79%
上海	95%	59%	91%	68%	78%
内蒙古	68%	53%	95%	89%	76%
北京	100%	0%	100%	100%	75%
黑龙江	95%	10%	95%	90%	73%
宁夏	100%	0%	94%	94%	72%
西藏	8%	0%	92%	0%	25%

图 18 各省份发布省市两级数据的持续性比例

3. 病例个案信息发布情况

除了关于某个地方疫情整体概况的统计数据，公众还想获知一些更为具体的、有关病人个体情况的数据，比如这些病人多大年龄？去过哪里？从哪里来？坐什么交通工具来？之前接触过什么人？有哪些病症？在哪家医院就诊？目前病情如何？和他接触过的人现在情况怎样？基于这些信息需求，一些省份开始发布更为详细的、有关确诊病例的个案信息（如以下截图）。

第35号确诊病例，男，8岁，系36号病例亲属，常住地湖北武汉。2020年1月22日与亲属一起乘CA4933次航班从武汉到南宁，转乘JD5588次航班前往三亚→1月23日02:00左右到三亚，乘酒店专车到三亚亚特兰蒂斯酒店入住，当天在酒店未外出→1月24日15:20左右前往亚特兰蒂斯水族馆游玩→1月25日17:10左右乘出租车陪其母亲前往301医院就诊，并在医院隔离观察→1月27日确诊，目前正在定点医院隔离治疗。

第36号确诊病例，女，17岁，系35号病例亲属，常住地湖北武汉。2020年1月22日与亲属一起乘CA4933次航班从武汉到南宁，转乘JD5588次航班前往三亚→1月23日02:00左右到三亚，乘酒店专车到三亚亚特兰蒂斯酒店入住，当天在酒店未外出→1月24日15:20左右前往亚特兰蒂斯水族馆游玩→1月25日17:10左右乘出租车陪其母亲前往301医院就诊，并在医院隔离观察→1月27日确诊，目前正在定点医院隔离治疗。

第37号确诊病例，女，3个月，常住地湖北孝感市。一家5人1月25日晚23:00左右自驾到达海口，当晚被安排在海口兵工大酒店入住→2020年1月26日10:00左右因发热被120救护车送至海口市人民医院住院隔离治疗→1月27日确诊，目前正在定点医院隔离治疗。

第38号确诊病例，男，37岁，常住地湖北武汉。2020年1月23日上午自驾从武昌出发前往海南→1月24日19:00左右进入海安新港安检，乘双泰26号轮渡前往海口，当晚到海口市维也纳国际酒店入住→1月25日早上在酒店用餐，11:00左右驾车到琼海市博鳌镇幸福海小区家中，晚上亲属来访→26日上午因身体不适，到琼海市人民医院就诊，14:30分左右和家人一起到医院对面的泰和酒店一楼餐厅吃饭，16:00左右住院隔离治疗→1月27日确诊，目前正在定点医院隔离治疗。

第39号确诊病例，女，64岁，常住地黑龙江。2019年11月30日乘火车从苏州前往海南→2019年12月2日到达东方市涛升国际小区住处→2020年1月9日参加海外国际旅行社组织的赴泰国旅行（东方同行11人）→1月15日乘SL8438次航班从泰国回到海口，乘汽车返回东方市八所镇→1月16-18日居家未外出→1月19日乘公交到海边游玩→1月20-25日居家未外出→1月26日乘15路公交车到东方市人民医院隔离观察→1月28日确诊，目前正在定点医院隔离治疗。

图 19 天津市卫健委官网发布的病例个案信息截图
(截图于 2020.1.29)

3.1 公开率

公开率是指各省份公布的确诊病例个案信息数占当地确诊病例总数的比例（公开率 = 确诊病例公布个案信息数 / 确诊病例总数），也就是说，在该省份已确诊的病例中，有多少比例的病例个案信息已经公布。图 20 展示了各省份公布病例个案信息的公开率。陕西和西藏公布了所有确诊病例的个案信息，达到 100%；贵州、吉林和天津公布了九成以上确诊病例的个案信息；而安徽、广东、湖北和江西公布确诊病例个案信息的比例最低（不足 0.5%）。

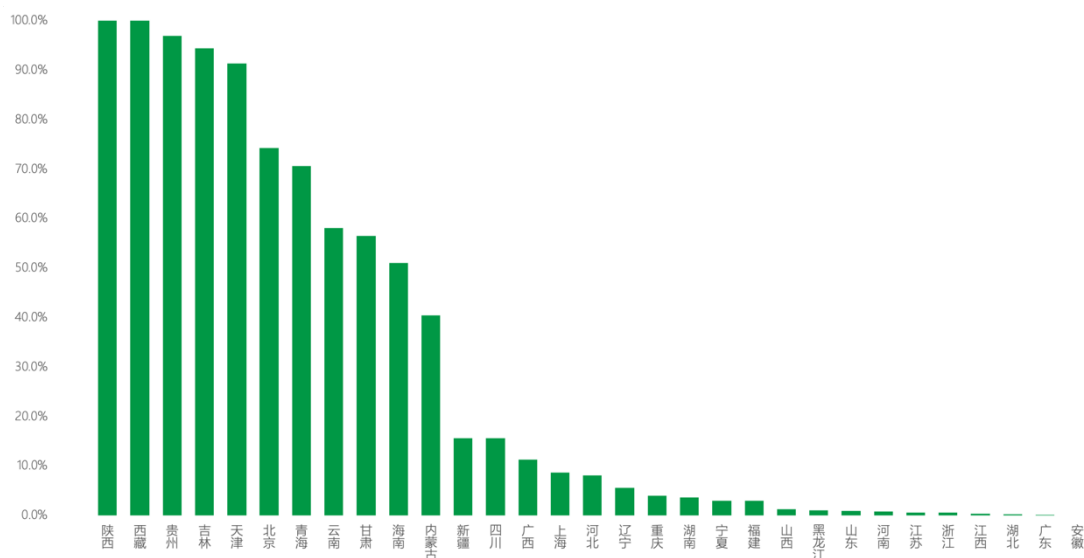


图 20 各省份确诊病例个案信息的公开率

3.2 精细度

精细度是指各省份发布的确诊病例个案信息的精细和具体程度。报告对各省份发布确诊病例个案信息的精细度进行了分析，包括基本信息、行为描述和诊疗情况三个方面，具体包含的信息内容、说明和举例见下表。

表 2 确诊病例个案信息发布内容

数据内容	说明	举例
基本信息	性别	“男，42岁”
	年龄	
	居住区域	“居于天津市河东区”
	居住小区	“怀德镇三合堡村”； “住址为双辽市北兴小区”
	工作性质	“在武汉华南海鲜市场从事售卖工作”
	与其他病例亲属关系	“该病例为病例1的母亲”
行为描述	重点地区接触史	“湖北武汉人，1月21日从武汉乘机到达腾冲”
	其他密切接触史	“系75号病例的密切接触者”
	其密切接触者情况	“该患者的密切接触者均在接受医学观察”
	交通工具概述	“乘坐机场巴士到达郑州机场”
	交通工具具体信息	“乘Z4178次列车到达郑州”
	发病期间活动场所	“到文昌龙楼镇石头公园游玩”
诊疗情况	发病时间	“1月14日发病”
	初次就诊时间	“1月20日到陕西省人民医院就诊”
	就诊医院具体名称	“先后就诊于湖北省人民医院、医大总医院”
	就诊医院模糊名称	“1月19日在本市发热门诊就诊”
	症状描述	“返回当晚出现发热”
	病情描述	“患者目前病情稳定”； “经市专家组确诊，为重症病例”
	当前状况	“被收治入院隔离治疗”； “于今日治愈出院”

图 21 展示了各省份发布确诊病例个案信息的精细度，颜色越深的方框表示该省份在其发布的确诊病例个案信息中包含了该项内容的比例越高。

总体上，在报告观测期内，大部分省份发布的确诊病例个案信息内容很少，从各省份的平均比例来看，发布确诊病例个案的性别和年龄这两项内容的比例相对较高，其次为发布当前状况的比例，而发病期间活动场所的比例最低。在确诊病例个案信息发布内容的精细度上，天津和陕西整体比例最高（分别为 52% 和 50%）。

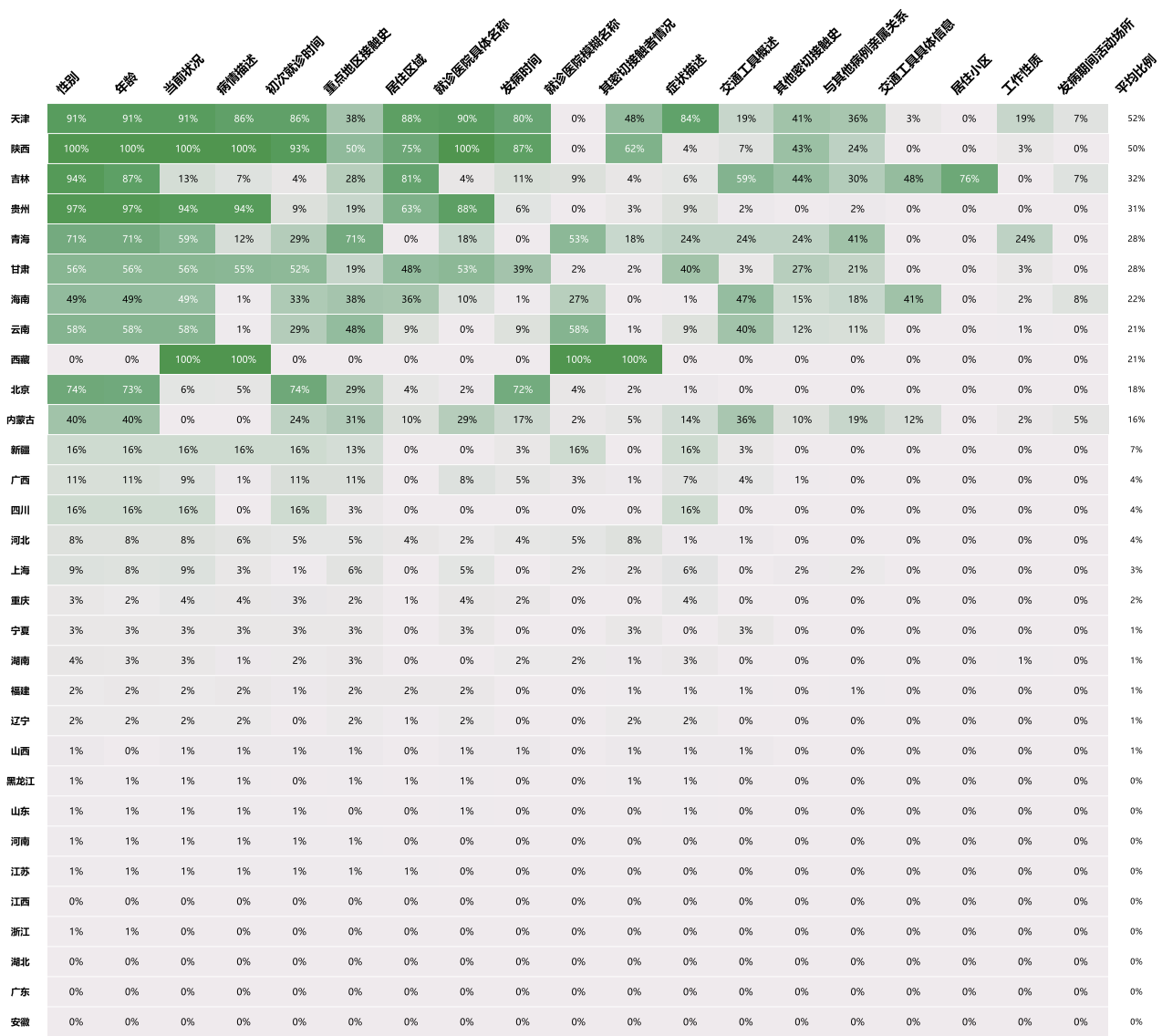


图 21 各省份发布确诊病例个案信息的精细度

图 22 展示了各省份发布的确诊病例个案信息的分布情况，图中方块越大，颜色愈深，表示该省份发布的确诊病例个案信息中包含了该项内容的比例越高。总体上，报告观测期内，各省份发布确诊病例的基本信息和诊疗情况的比例高于发布行为描述的比例。



图 22 各省份发布确诊病例个案信息的分布情况

3.2.1 居住区域 / 小区信息

报告发现，关于确诊病例的居住地点，有些省份在发布的确诊病例个案信息中包括了病例所居住的行政区域（省、市、区县等），而有些省份则发布了更为具体的病例所居住的小区或村等信息。图 23 展示了各省份在其发布的确诊病例个案信息中包含了居住区域和居住小区信息的比例。在报告观测期内，天津发布了近九成确诊病例的居住区域信息（88.4%），吉林发布了八成以上确诊病例的居住区域信息（81.5%）；但仅有吉林发布了超过七成确诊病例的居住小区信息（75.9%）。

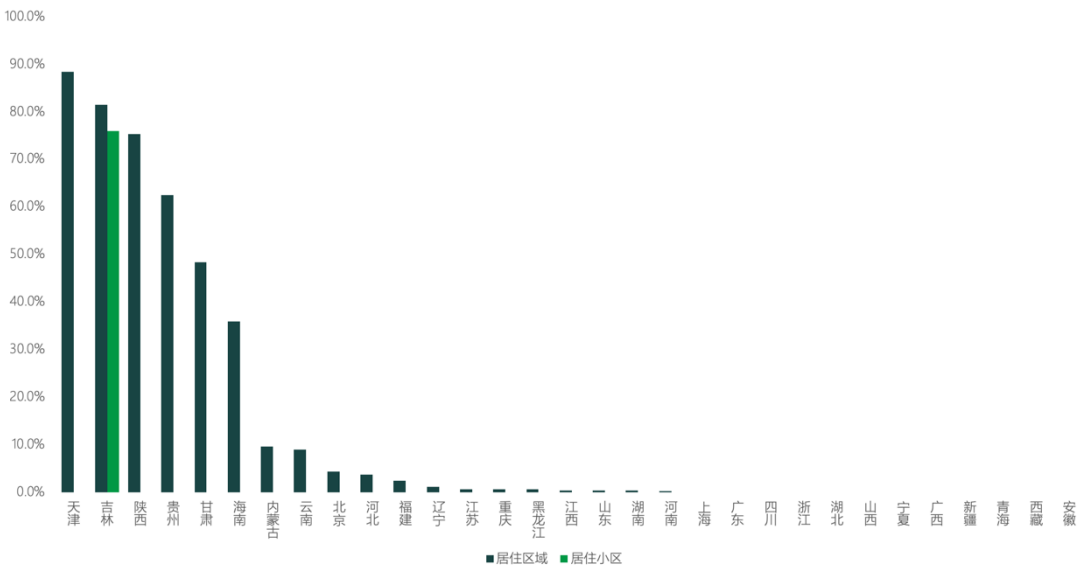


图 23 病例个案信息中包含居住区域 / 居住小区的比例

3.2.2 重点地区接触史 / 其他密切接触史信息

接触史信息包括重点地区接触史和其他密切接触史信息。重点地区接触史信息是指病例曾在疫情重点地区生活、暂居或途经过；其他密切接触史信息是指病例与已确诊病例有过密切接触而感染。图 24 展示了各省份发布确诊病例接触史信息的比例。报告观测期内，青海发布了七成病例的重点地区接触史信息和两成以上病例的其他密切接触史信息（70.6% 和 23.5%），陕西发布了五成病例的重点地区接触史信息和四成以上病例的其他密切接触史信息（50.0% 和 43.4%），天津发布了五成病例的重点地区接触史信息和四成以上病例的其他密切接触史信息（50.0% 和 43.4%），陕西发布了五成病例的重点地区接触史信息和四成以上病例的其他密切接触史信息（分别为 50.0% 和 43.4%）。

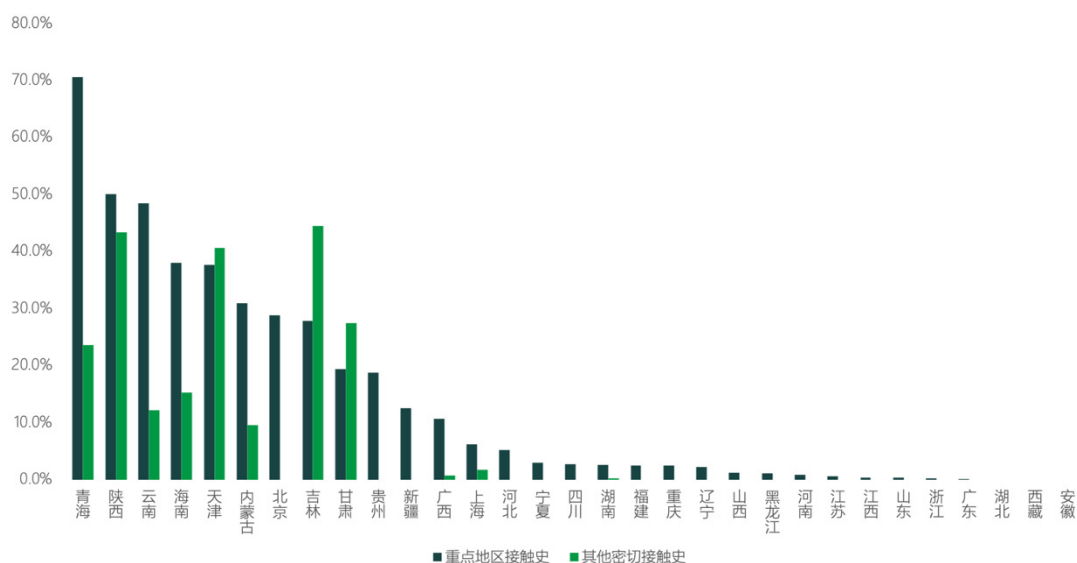


图 24 病例个案信息中包含重点地区 / 其他密切接触史信息的比例

有些确诊病例与其他已确诊病例之间存在亲属关系，部分省份也发布了该类别信息，例如在病例情况中提到“该病例为病例 1 的母亲”等。图 25 展示了各省份发布的确诊病例与其他已确诊病例亲属关系信息的比例。报告观测期内，青海和天津在所发布的病例中包含了与其他病例亲属关系信息的比例相对最高（分别为 41.2% 和 36.2%）。

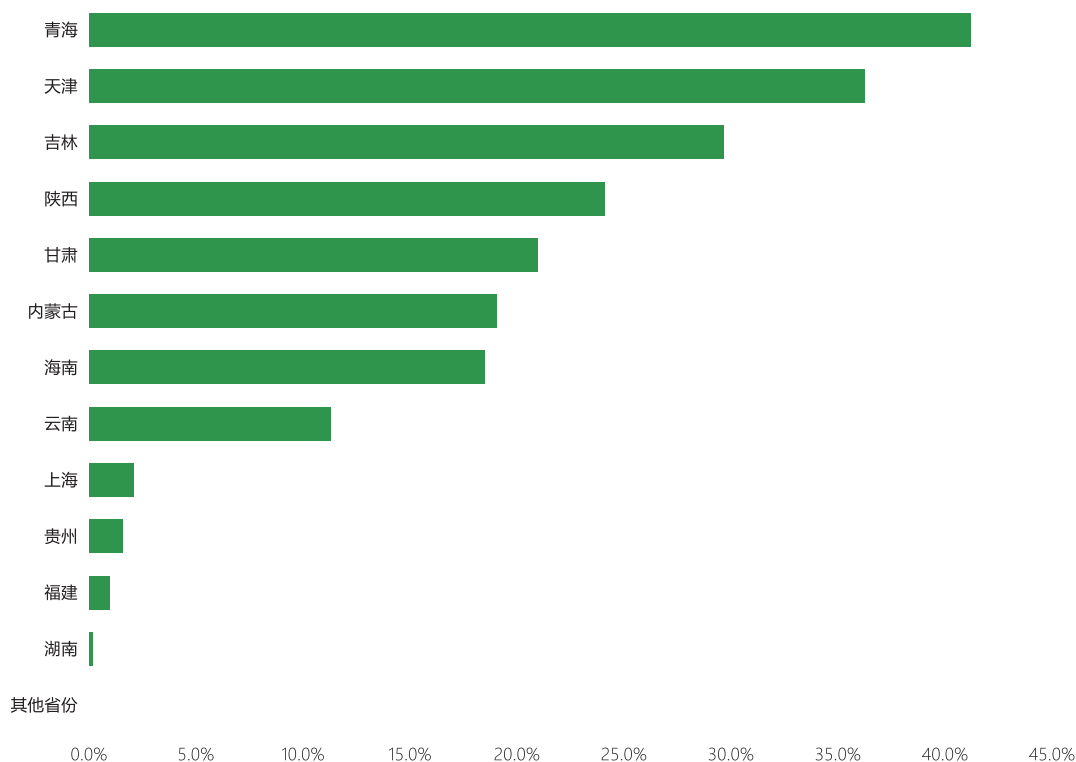


图 25 病例个案信息中包含与其他病例亲属关系信息的比例

3.2.3 交通工具概述 / 交通工具具体信息

关于确诊病例在发病期间乘坐过的交通工具，有些省份在发布的病例个案信息中对交通工具进行了概述，例如乘坐高铁、飞机、公交车等；而有些省份还发布了交通工具的具体信息，即飞机航班、高铁车次、公交车车牌等。图 26 展示了各省份在确诊病例个案信息中发布了交通工具概述 / 交通工具具体信息的比例。报告观测期内，吉林发布了近六成确诊病例（59.3%）的交通工具概述和近五成确诊病例（48.1%）的交通工具具体信息，海南公布了近五成确诊病例（46.7%）的交通工具概述和超过四成确诊病例（41.3%）的交通工具具体信息。

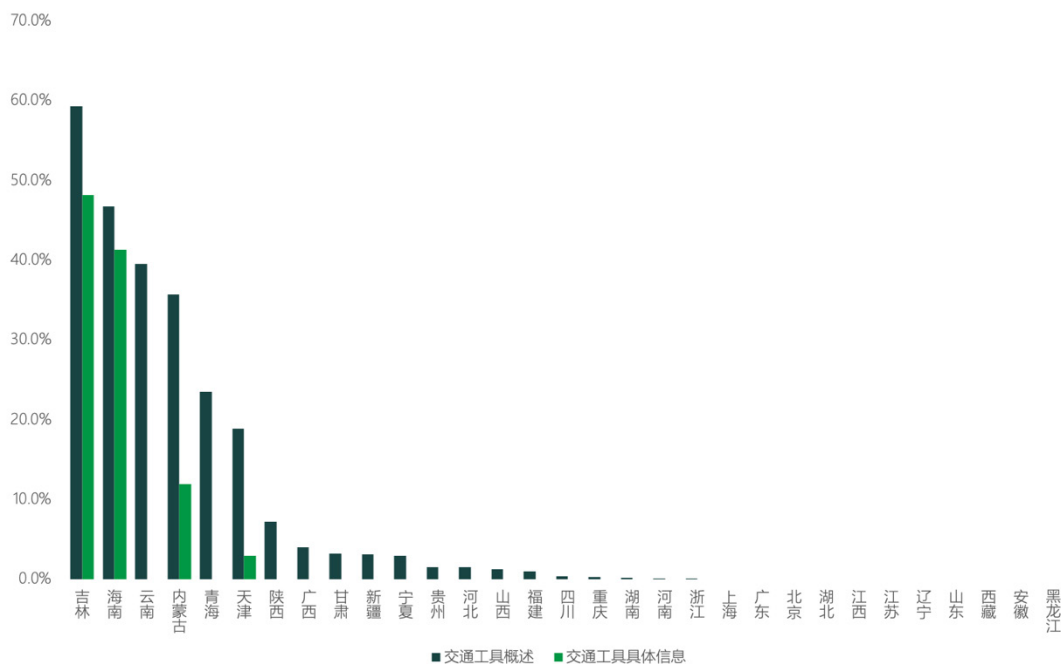


图 26 病例个案信息中包含交通工具概述 / 具体信息的比例

3.2.4 发病期间活动场所信息

关于确诊病例的发病期间活动场所，有些省份在病例个案信息中发布了具体的确诊病例在发病期间曾逗留过的餐厅、酒店和公园名称等信息。图 27 展示了各省份在确诊病例个案信息中发布病例活动场所信息的比例。报告观测期内，仅有海南、吉林、天津、内蒙古发布过部分确诊病例发病期间的活动场所信息，且发布比例均不足一成。

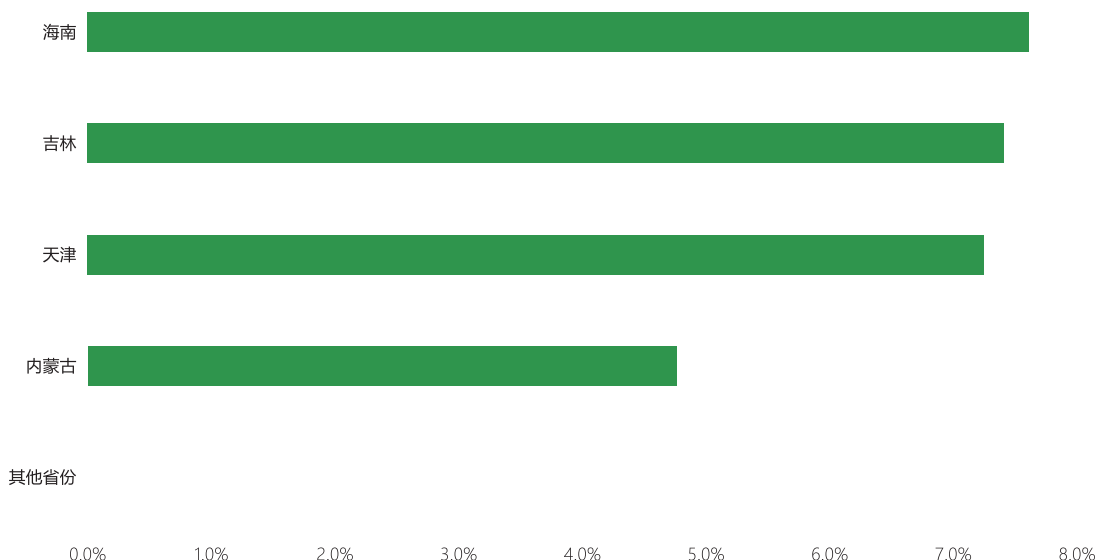


图 27 病例个案信息中包含发病期间活动场所信息的比例

3.2.5 就诊医院模糊名称 / 就诊医院具体名称

关于确诊病例的就诊医院，有些省份在病例个案信息中发布了模糊的就诊医院名称，例如“定点医院”“发热门诊”“医疗机构”“救治中心”等；而有些省份还发布了具体的就诊医院名称。图 28 展示了各省份在确诊病例个案信息中发布了就诊医院模糊名称 / 具体名称的发布比例。在报告观测期内，陕西公布了所有确诊病例就诊医院的具体名称，天津和贵州公布了近九成确诊病例就诊医院的具体名称（分别为 89.9% 和 87.5%）。

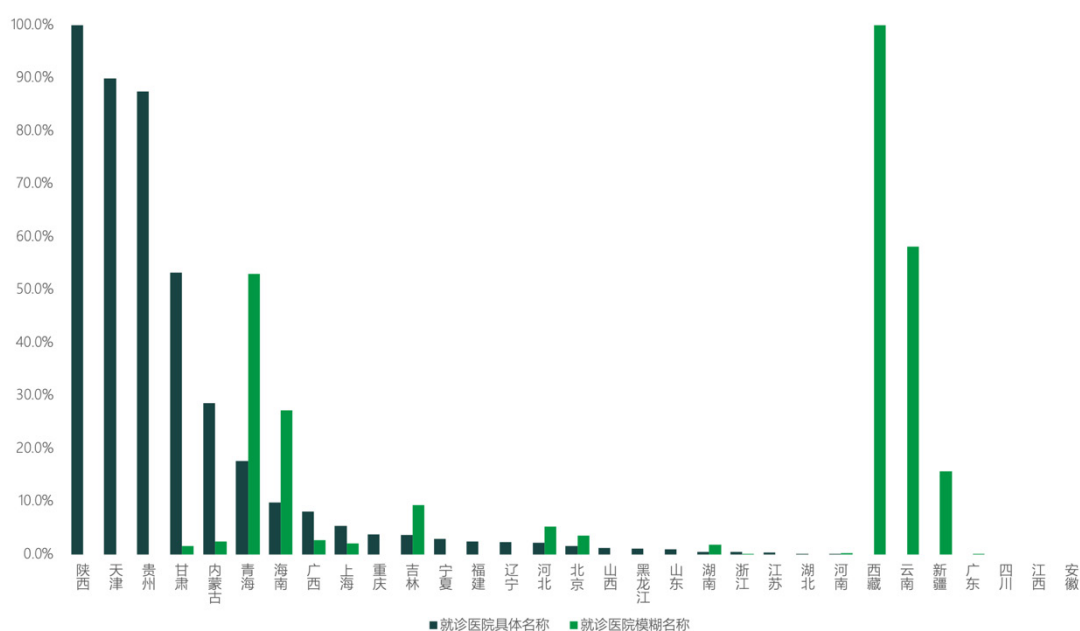


图 28 病例个案信息中包含就诊医院模糊名称 / 具体名称的比例

3.3 隐私保护

报告也发现有少数省份发布了过于详细的、可直接锁定到个人的数据，如病例的私家车车牌号等。还有些省份同时发布了某个患者的姓氏、性别、年龄、居住的街镇和就诊医院等信息，这些数据如分别发布，风险相对可控，但同时发布且可关联到同一个患者时，很容易被用来锁定到患者本人，暴露病人的隐私（如以下截图）。

新增10例患者信息如下：

患者1：顾某某，男，42岁，外地来黔人员，现住黔西南州义龙新区顶效街道。2月3日被诊断为确诊病例，目前在黔西南州人民医院隔离治疗，病情平稳。

患者2：何某某，男，27岁，外地来黔人员，现住黔南州贵定县云雾镇。2月3日被诊断为确诊病例，目前在贵定县人民医院隔离治疗，病情平稳。

患者3：陈某某，男，30岁，外地来黔人员，现住遵义市仁怀市九仓镇。2月3日被诊断为确诊病例，目前在遵义市第一人民医院隔离治疗，病情平稳。

患者4：朱某某，女，61岁，本省常住人口，现居于贵阳市南明区太慈社区。2月3日被诊断为确诊病例，目前在贵州医科大学附属医院隔离治疗，重症。

患者5：朱某某，女，53岁，外地来黔人员，现住贵阳市南明区太慈社区。2月3日被诊断为确诊病例，目前在贵州医科大学附属医院隔离治疗，病情平稳。

患者6：江某某，女，1个月，外地来黔人员，现住贵阳市南明区花果社区。2月3日被诊断为确诊病例，目前在贵州省人民医院隔离治疗，病情平稳。

患者7：饶某，女，27岁，本省常住人口，现居于毕节市赫章县白果街道。2月3日被诊断为确诊病例，目前在毕节市第三人民医院隔离治疗，病情平稳。

患者8：文某，男，29岁，外地来黔人员，现住毕节市赫章县妈姑镇。2月3日被诊断为确诊病例，目前在毕节市第三人民医院隔离治疗，病情平稳。

患者9：陈某某，女，56岁，本省常住人口，现居于毕节市大方县红旗街道。2月3日被诊断为确诊病例，目前在大方县人民医院隔离治疗，病情平稳。

患者10：陈某，女，20岁，本省常住人口，现居于毕节市大方县顺德街道。2月3日被诊断为确诊病例，目前在大方县人民医院隔离治疗，病情平稳。

**图 29 贵州省卫健委官网发布的病例个案信息截图
(截图于 2020.2.4)**

为了准确查找与郑某某有接触的人员，现发布通报：

郑某某离开武汉之后的具体行程轨迹：

1.1月16日在武汉市汉口站与其丈夫一同乘火车回到长春，列车号T182次，座位号7车30号座、31号座，汉口开车时间14时57分，1月17日15时22分到达长春站。

2.到达长春站后，在长春北站凯旋路客运站乘客车准备到公主岭市怀德镇下车，车牌号吉C3D444。由于车上乘客仅为6人，后转乘其他客车，在长春市绿园区春城大街百家利商场门前上车，车牌号吉C39977，到怀德镇三十里堡站下车。

3.16:47分左右，乘私家车，车牌号吉JB2998，到达怀德镇三合堡村亲属家居住。18日在家休息。

4.1月19日乘私家车，车牌号吉JB2998，于11时左右到秦家屯戡子街刚子饭店就餐，随后，原车返回怀德镇三合堡村。

5.1月20日，郑某某与其丈夫乘客车，怀德镇三合堡村上车，车牌号吉CK0557，于上午9时左右到公主岭市中心医院神经内科6楼630病房看望病人，10时左右步行到公主岭市温州商城，在温州城裙房短暂逗留，在商城附近二区西门对面鑫诚源家常菜饭店用餐。

6.11时53分左右到公主岭市客运站买票，12时左右乘客车，车牌号吉CK0557，回到怀德镇三合堡村其女婿家居住。

7.1月21日下午，郑某某与其丈夫乘私家车，车牌号吉JB2998，到秦家屯镇南平村5队其亲属家居住。

8.22日到25日在家休息。26日因发热，到公主岭市中心医院治疗。

**图 30 吉林省卫健委官网发布的病例个案信息截图
(截图于 2020.1.29)**

表 3 展示了报告观测期内所发现的存在病例隐私泄露风险的省份、病例数量及相关的信息内容。

表 3 病例个案信息发布中的隐私泄露风险情况

省/直辖市	存在隐私泄露风险的病例数量	可关联到同一病例的、存在隐私泄露风险的信息内容
贵州	19	姓氏、性别、年龄、所居街道/镇/村、就诊医院
天津	8	姓氏、性别、年龄、所居镇、工作单位
吉林	5	性别、年龄、所居镇/村/小区、车牌号
云南	3	性别、年龄、所居村
海南	3	姓氏、性别、年龄、工作单位、职务、所居镇/村
湖南	1	姓氏、性别、年龄、所居街道、工作性质

讨论与调查

报告还针对疫情数据发布的透明和细致程度，从互联网上采集了专家评论、媒体观点和网友点评，并进行分析和梳理，以下列举了其中具有代表性的观点。

1. 数据发布应公开、及时、详细的理由

1.1 保障公众知情权，提升政府公信力

“政府有关部门公布这些信息也是履行政府及时、准确和全面公布疫情信息的政府职责，是落实民众知情权的具体手段。”（法学专家观点）

“现在是一个高度信息化的时代，人人都有麦克风，人人都是自媒体。信息不公开或者迟公开，只会助长各种小道消息满天飞，自然也会掺杂一些不实信息和谣言，从而对政府的公信力造成伤害。”（媒体观点）

“某些地区公布疫情确诊病例信息含糊，仅公布所在城区，未公布接触重点地区人员的详细时间轨迹以及返回当地后至确诊期间的详细行动轨迹，各种民间消息真真假假群众难以分辨，政府公信力岌岌可危。”（网友观点）

1.2 消除坊间谣言，减少社会恐慌情绪

“基于公开透明的原则，回应近日广大网民普遍提出的一些疑问和困惑，以减少不必要的社会恐慌。”（政府部门观点）

“及时、公开、详细的信息，满足了公众的知情权，也消弭了滋生谣言的空间。”（专家观点）

“发布详细的病例信息，也能体现政府有所作为，已经做了详细的疫情调查工作，越详细的数据发布越能让大家安心。”（网友观点）

“与其道听途说，让民众猜猜猜，不如公开信息，让大家坦然面对，做好防护，克服恐慌。”（网友观点）

“公开不太会导致恐慌，反而因为能采取应对措施，感觉更安心了。”（网友观点）

“面对一场充满诸多不确定性的严重疫情，公众难免会感到焦虑，甚至恐慌。而消除公众焦虑和恐慌的最有效办法，就是把公众最想知道的信息及时公布出来。”（媒体观点）

“官方版本替代坊间流传版本，让市民获知可靠的信息。”（媒体观点）

1.3 助力精准防控，确定重点区域与人群

“越是最紧要的时候，越要让防控更加精准、更加科学、更加有力。最大限度地公开疫情信息，就是精准防控的关键。”（媒体观点）

“一方面有利于所有人准确获悉、采取针对性的防范措施，另一方面，便于社区、民众有针对性地开展防控工作，做好源头防控，遏制疫情扩散蔓延。”（媒体观点）

“患者轨迹是流行病学调查中的一个重要环节，因地制宜、因人而异地公布这些信息可以让社区、普通民众确定疾病传播的高风险区域，进行更有针对性的防控。”（学者观点）

“因为精确到分钟、人员关系清晰，既有利于纾解恐慌情绪，又方便找到密切接触者，天津官方公开的确诊病例行动轨迹近来一直备受好评。”（媒体观点）

“有了真实详尽的疫情图，更多的市民心中有数，就能更加有效地配合政府防控工作。”（媒体观点）

1.4 提高风险意识，引导市民加强自我防护

“公布新确诊病人在出现病征期间曾经逗留的场所和活动轨迹，某种程度上就是一种强力安全提示，以帮助居民进一步消除麻痹大意心理。”（媒体观点）

“在基于保护隐私的前提下公开更多的信息，有助于公众知情并自我防范。”（政府相关部门负责人观点）

“市民不满足于知道有多少确诊病例、疑似病例，不满足于一个个枯燥的数字，还想知道他们发病前的活动轨迹等更详尽的信息，以便采取更合理的应对措施。”（媒体观点）

1.5 便于居民自查，提醒个人防护观察

“人们急于知道，自己究竟是否曾经跟确诊者身处同一空间，是否是潜在的接触者，是否有被传染的可能，是否需要主动居家隔离 14 天。”（网友观点）

“公布轨迹的意义，就是清楚明白地提醒公众，哪里曾有危险、是否去过，若去过就多注意监测和隔离。”（媒体观点）

“详细的轨迹公布，能更好更细致地找到相关人员。如果市民与公布的确诊病例活动轨迹有实际交叉，就可以注意早期诊断，早期隔离。一方面主动上报，一方面居家隔离，同时检测体温并做记录。”（专家观点）

“提醒可能有过交集或可能会有交集的自动在家隔离，才是务实防控。”（网友观点）

2. 数据发布应谨慎、适度、有用的理由

2.1 泄露患者隐私，影响其正常生活

“患者信息公布得越详细，对其身份的认定就越容易，从而造成其个人隐私的泄露，对他们的生活造成不良影响。”（媒体观点）

“民众对于数据信息的要求是无止境的，公布了病例轨迹，接下来就要求公布病例活动的具体时间，还要求公布病例的住址、病例的密切接触者等信息，这会让患者隐私保护变得越来越困难。”（媒体观点）

2.2 引发无端猜测，加剧社会恐慌

“信息公开过多会引发很多无端猜测，网上有人发热居家隔离后被视作瘟神遭到各式骚扰，还有些与确诊患者同小区的居民外出办事时，也被当成可疑人物对待。”（网友观点）

“公布涉及区域和场所，对寻找密切接触者的意义很大，但对于非密切接触者，也可能会引发无端的猜测。”（专家观点）

“如果把确诊患者所有活动过的地点，都对大家进行公示的话，在某些省份、某些场合可能会引起一定的恐慌。”（专家观点）

2.3 缺乏实际效用，部分公共场合无法避免

“一些公共场所无法避免，信息公开没有必要。公布信息中不乏有某地铁站等公众场所，但最主要还是主管部门做好全面的消毒工作，每位市民都要做好个人防护。”（专家观点）

“轨迹公布的有些公共场合是难以避免要去的，相关部门已经做好了防疫措施，不明白公布这些省份的意义何在？”（网友观点）

2.4 造成麻痹思想，降低居民防范意识

“很多人居然想着活动区域不在名单上，与公布的病例轨迹没有交集，就可以出去逛了。”（网友观点）

3. 调查结果：数据应该发布到多细？

针对“疫情数据应该发布到多细”这一问题，报告进行了初步的网络调查，共收到反馈问卷 576 份，调查结果如下。

3.1 发布病例个案数据

整体上，关于是否应该发布每一个病例的基本情况和行踪轨迹，但不公布其姓名等个人信息，选择支持的比例高达 93%（如图 31）。

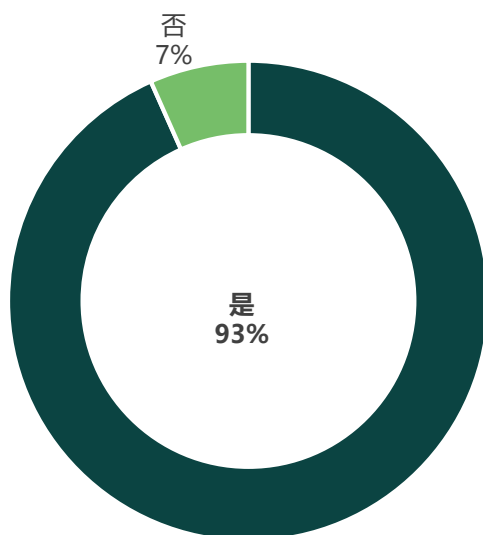


图 31 关于发布病例个案数据

3.2 发布病人的居住地点

具体而言，关于病人的居住地点应该发布到多细？调查结果如图 32 所示，当要求发布的数据越来越接近病人隐私的时候，选择反对的比例逐渐增高。对于是否应该发布病人居住的小区名称，选择支持的比例高达 97%；对于是否应该发布病人居住的小区名称和楼号，选择支持的比例降到了 64%；而对于是否应该发布病人居住的小区名称、具体楼号和室号时，选择支持的比例仅为 19%。

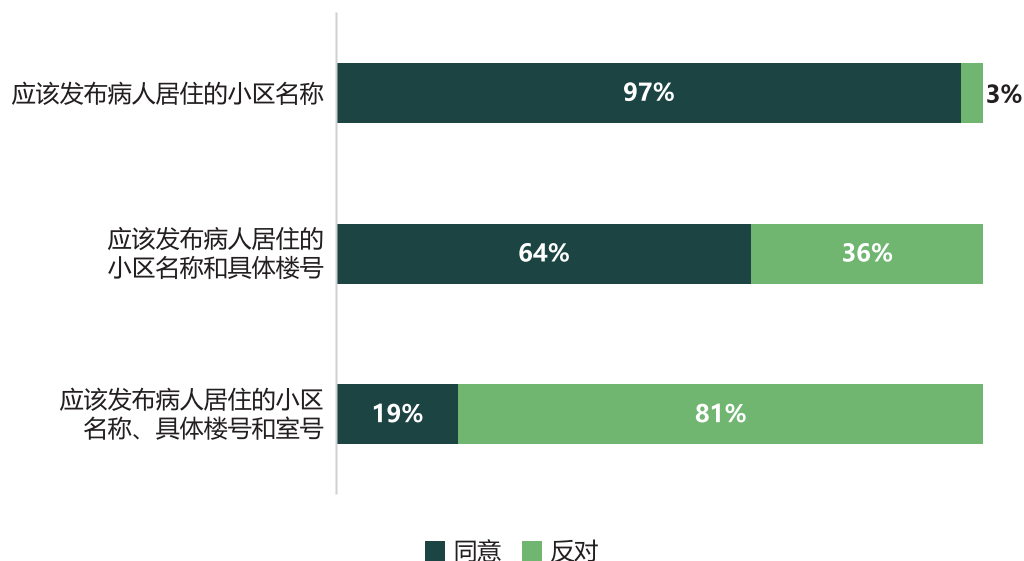


图 32 关于发布病人的居住地点

3.3 发布病人乘坐过的交通工具数据

关于发布病人乘坐的交通工具的数据应该多细？调查结果同样显示（如图 33），当要求发布的数据越容易泄漏病人隐私时，选择反对的比例越高。对于是否应该发布病人乘坐过的飞机航班（火车车次）和日期，选择支持的比例 98%；对于是否应该发布病人乘坐过的飞机航班（火车车次）、日期和座位号，选择支持的比例降到了 70%；而对于是否应该发布病人自驾车的车牌号这样的隐私数据，选择支持的比例仅为 15%。

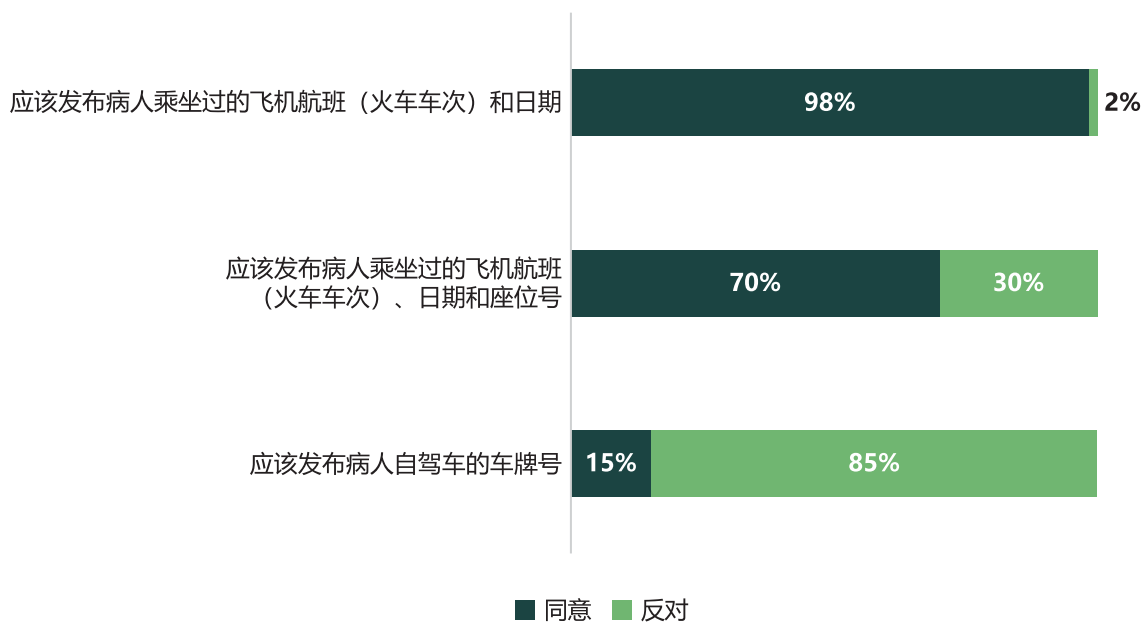


图 33 关于发布病人乘坐过的交通工具

3.4 发布病人的活动场所

关于病人的活动场所应该发布到多细？调查结果如图 34 所示，对于是否应该发布病人去过或来源的城市名称，选择支持的比例为 98%；对于是否应该发布病人住过的酒店、吃过饭的餐厅、去过的商场或其他场所，以及是否应该发布病人去就诊过的医院，选择支持的比例略降到 93%。

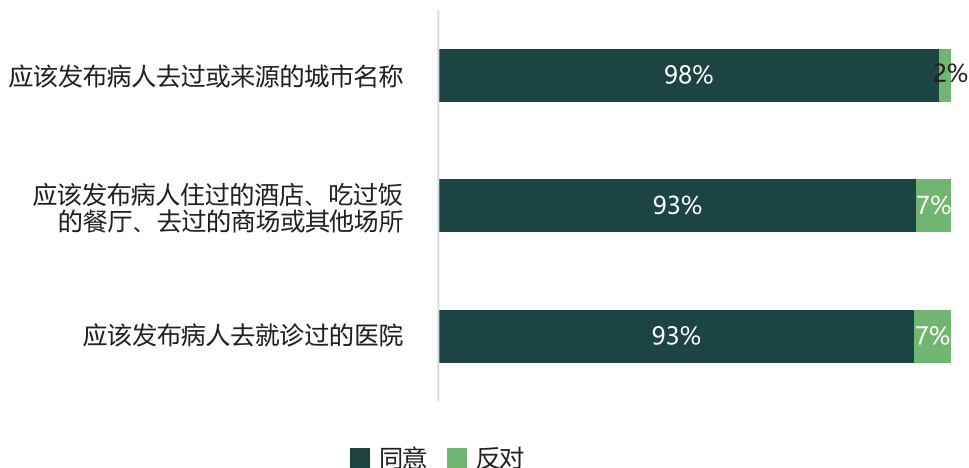


图 34 关于发布病人的活动场所





结论与建议

1. 总体情况

在整体发布情况上，各省份在数据发布的时效性、易得性、可读性等方面表现参差不齐。在当地疫情统计数据发布上，没有一个省份连续发布了每项数据，各省份发布的数据或是不全面，或是不连续。在病例个案信息的发布上，各省份发布的个案数量和信息内容各有差异，没有一个省份发布了所有确诊病例个案的每项信息。

报告发现，各省份疫情统计数据的发布水平整体高于病例个案信息的发布水平。并且，随着时间的推移，各省份发布疫情数据的水平也在不断提升和完善，数据的全面、持续和细致程度逐步提高。但总体上，公众对于疫情数据的需求尚未得到数据供给的充分满足。

在各个维度上疫情数据发布水平相对领先的省份的分布情况如图 35 和表 4 所示。在图中，某个省份的颜色越深，代表该省份发布水平领先的维度越多，例如吉林、陕西、天津、上海、山东、重庆等省份，都在两到三个维度上相对领先。

然而，作为国内最先发现病例，疫情也最严重的省份，湖北省卫健委官网直到 1 月 21 日才开始发布疫情数据，报告观测期内，几乎在发布内容和发布方式的各个维度上都处于靠后位置。

同时，其他新冠肺炎累计确诊人数较多的省份，如广东、河南、浙江、湖南、安徽、江西等省份也都未进入疫情数据发布水平相对领先的名单。一个地方的疫情严重程度与疫情数据发布水平之间是否存在相关关系或因果关系？还有待进一步开展研究。



图 35 各维度上数据发布水平相对领先的省份

开放数据

表 4 各维度上数据发布水平相对领先的省份

一级维度	二级维度	疫情数据发布水平 相对领先的省份
整体发布情况	时效性	重庆、天津、上海
	易得性	15个省份并列
	可读性	甘肃、吉林、四川
疫情统计数据发布情况	全面性	重庆、吉林、山西
	持续性	上海、山东、福建
	覆盖层级	海南、山东、福建
病例个案信息发布情况	公开率	陕西、西藏、贵州
	精细度	陕西、天津、吉林

2. 问题与建议

报告发现，虽然我国有关传染病防治和突发公共卫生事件应急的法律法规对发布疫情信息已提出了“及时、全面、准确”的要求，但从目前各地实际发布水平参差不齐、缺乏规范的现状来看，我国还需要进一步完善相关法律法规，并制定更为具体的、具有可操作性的实施细则和数据标准，从而使疫情数据开放工作更加依法合规和统一有序。具体而言，可从以下几方面对疫情数据发布的内容和形式作出规定。

2.1 发布时效：及时和持续

在移动互联和社交媒体时代，信息传播的速度和广度都远超以往，政府如未能在第一时间发布权威真实的疫情数据，各种真真假假的小道消息将会迅速传播，增加社会恐慌情绪。同时，缺少必要的提醒和警示，也会降低公众的自我防护意识，造成疫情扩散。

因此，在疫情期间，政府部门应充分保障公众知情权，及时、持续、充分地发布疫情数据，这将有助于压缩谣言空间，消除社会恐慌情绪，提高公众的自我防护意识，防止疫情大规模扩散，并提升政府公信力。

2.2 发布内容：全面和精细

进入大数据时代，公众在疫情期间的信息需求也出现了新的变化，他们不仅需要获得有关疫情的政策文件、工作动态和防护知识等信息，还希望能获得真实的、全面的、详细的数据。

面对这一需求，政府首先需要发布有关疫情真实发展情况的统计数据，覆盖疫情防控和治疗的全周期，包括接受观察、疑似、排除观察、确诊、病情平稳、重症、危重、出院、死亡等各个关键环节。

然而，统计数据是对原始数据进行加工归总后形成的结果，数据的利用价值有限。

例如，面对这样一条统计数据，“已发现确诊病例 326 例。确诊病例中，男性 168 例，女性 158 例；年龄最大 88 岁，最小 7 月龄”，我们无法得出在 7 个月到 88 岁这个年龄区内，病人主要集中在哪几个年龄段？在某一个年龄段内是男性多还是女性多？

要回答这些问题，就需要政府发布更为详细的有关病例个案的信息，包括基本信息、行为描述信息和诊疗情况信息等。这些一手的、未经加工归总的、更为原始和精细的数据，比统计数据具有更大的信息量和利用潜力。

此外，省级网站除了发布省本级的疫情数据外，还应发布下辖各地级行政区域的数据，便于公众在一个平台上集中便捷地获取整个省份的疫情数据；与此同时，也应允许和鼓励下辖的地级政府在本地平台上发布疫情数据，使公众有更多渠道获知与自己居住地区相关的疫情数据。

2.3 发布标准：规范和完整

报告发现，不同省份在数据名称的表述上略有差异。例如，有的省份将“病情平稳”作为一种病例类型，来区分于危重病例；有些省份则将“病情平稳”作为一种病情状态来描述轻症和重症病例，然而，在发布了这些数据的页面上，都未能找到关于这些专业术语的解释说明。普通公众并不是医学或公共卫生专家，政府在发布疫情数据时用词不统一，又未提供说明，容易造成公众误读误用这些数据。因此，政府在发布疫情数据时应统一数据标准，并配备详细的备注或说明文件，帮助公众准确无误地解读数据。

报告还发现，不同省份在发布数据的完整性上也有待改善。例如，在累计重症 / 危重病例这项数据上，有些省份在当天并无重症病例出现的情况下，仍坚持发布该项数据，只是将数据的数值发布为“0”；而有些省份则未能每天发布该项数据，使公众无法区分当天到底是“没有报”该项数据，还是该项数据的数值为0，造成数据缺失。因此，政府在发布疫情数据时还应保持完整性和连续性。

2.4 发布形式：易得和可读

报告发现，政府目前发布的疫情数据通常以碎片化的方式散落在不同时间发布的、位于不同页面上的、以不同名称出现的通告里，不便与公众获取。同时，这些数据普遍呈现为文字夹杂数据的形态，不便于公众阅读和分析。结果造成，虽然政府从其自身视角出发，以他们惯用的方式发布了疫情数据，但老百姓却没有获得感。因此，政府发布疫情数据应保持其易得性和可读性，方便公众快速发现和获得数据，并通过结构化表格和可视化图形的方式，帮助公众阅读和理解数据。

在大数据时代，政府在“发布数据”的同时，还应进一步向“开放数据”迈进，便于社会对数据进行开发利用。目前许多省份虽然发布了疫情数据表格，但多是以图片格式提供的，这种格式虽然便于“阅读”，但还不易于“利用”，如要对这些数据进行分析或开发，需要先对图片中的数值进行识别，存储为电脑可以读取和处理的数据格式，如 excel、csv 等格式。

2.5 发布底线：法治和温度

政府部门在发布疫情数据时，既要充分保障公众的知情权，又要严格保护患者的隐私，平衡好两者之间的关系至关重要。特别要把握好数据的精细度和关联度，使数据既能起到提醒公众加强自我防护，降低恐慌情绪的作用，又不能被用于锁定到患者本人，造成歧视和伤害。应坚持运用法治思维和法治方式发布疫情数据，守住法治的底线。

每一条病例数据的背后都是活生生的人，因此政府发布的每一个数，都应带有人的温度。发布的数据太粗太慢，无法温暖人；发布的数据过细过火，则会烫伤人。政府防控疫情，是为了让每个人都健康平安，保持体温 37 度，不低也不高；政府发布疫情数据，也要控制好火候，让数据的温度始终保持在 37 度，不欠也不过。

附录

调查问卷：

您认为疫情数据应该发布到多细？

1. 应该发布病人居住的小区名称（单选）

A. 同意 B. 反对

2. 应该发布病人居住的小区名称和具体楼号（单选）

A. 同意 B. 反对

3. 应该发布病人居住的小区名称、具体楼号和室号（单选）

A. 同意 B. 反对

4. 应该发布病人去过或来源的城市名称（单选）

A. 同意 B. 反对

5. 应该发布病人乘坐过的飞机航班（火车车次）和日期（单选）

A. 同意 B. 反对

6. 应该发布病人乘坐过的飞机航班（火车车次）、日期和座位号（单选）

A. 同意 B. 反对

7. 应该发布病人自驾车的车牌号（单选）

A. 同意 B. 反对

8. 应该发布病人住过的酒店、吃过饭的餐厅、去过的商场或其他场所（单选）

A. 同意 B. 反对

9. 应该发布病人去就诊过的医院（单选）

A. 同意 B. 反对

10. 应该发布每一个病例的基本情况和行踪轨迹，但不公布其姓名等个人信息（单选）

A. 同意 B. 反对



联合发布方：



DMG Lab
Fudan University
复旦大学 | 数字与移动治理实验室



DATAMATE数据工作室

团队成员：

郑磊	吕文增	韩笑	陈瑶	邓婕	王近斐
张宏	侯铖铖	张忻璐	杜雨萱	赵慧	郑银银
刘新萍	周业光	华蕊	李珍		

上海市杨浦区邯郸路220号复旦大学智库楼404室

dmglab@fudan.edu.cn

🌐 www.dmg.fudan.edu.cn



关注
复旦DMG
公众号