

2022年12月14日

基于搜索指数还原疫情水平——宏观量化策论之一

宏观研究团队

——宏观经济点评

何宁（分析师）

hening@kysec.cn

证书编号：S0790522110002

● 防疫持续优化

11月以来我国防疫优化稳步迈进新阶段。11月11日“二十条”出台后，各地“因城施策”边际优化防疫政策，截至12月9日我们统计已有19个省市落实“新十条”。在“缩小核酸检测范围、减少频次”、“非必要不核酸”等导向下，以及多个省市取消常态化核酸，如何更好跟踪各地疫情水平以及判断第一轮冲击达峰情况？

● 搜索指数映射疫情水平

表观疫情数据的传导链条为新冠→居民→卫健委。居民在感染或认为感染概率较高时，其可能发生的行为包括搜索新冠本体、新冠体征以及购买药品等。若能有效观测到前半段传导，理论上或可领先于卫健委的表观疫情数据。

陈沁博士在《各城市首轮感染高峰期预测》提出以发烧指数衡量疫情感染情况，我们将其扩展为三类：搜索新冠本体、新冠体征以及购买药品，具体为“新冠”、“发烧”、“莲花清瘟”、“布洛芬”等关键词。

从与新增确诊的相关性来看，药品搜索指数>疫情搜索指数>体征搜索指数。有趣的是，“莲花清瘟”在三类关键词中相关性最好，附带的“莲花清瘟”同样具备映射疫情水平的作用。“二十条”发布后，莲花清瘟指数迅速上升而后下行，与疫情数据有小幅背离，在于防疫优化推进时药品包含了对未来疫情扩散的预期、因而提前上升；

发烧指数在中国的有效性似乎并没有那么高，这可能与我国奥密克戎变异路径有关。2022年上半年主要是ba.2和ba.2.2变异株、下半年为BA.5和BF.7，前两者感染后临床表现相对温和、无症状感染者占比高，但BF.7感染症状包括高烧、持续性咳嗽、身体疼痛不适、头痛、喉咙痛等，与新冠奥密克戎早期亚型变异株感染后的症状类似。因此，高烧指数无法描绘前三个季度的疫情水平，但11月以来有效性显著提升。

搜索指数具备高频、领先性、抗干扰等特点。我们将时间窗口选择为2022年1月1日-2022年12月7日，并计算各搜索指数与确诊病例的相关性。从全国来看，莲花清瘟指数、新冠指数、发烧指数与确诊病例的相关性分别为0.70、0.63、0.03；莲花清瘟指数的领先期主要分布在5-10日、平均为7.1日，新冠指数的领先期主要分布在2-8日、平均为6.1日。

● 利用搜索指数拟合疫情水平

“小步走、不停步”的防疫优化已初步迎来第一波冲击，我们尝试利用搜索指数拟合疫情水平。回归统计数据显示，20个地区调整R方超过0.5，表明该方法具有一定解释力度。

从时间序列来看，本轮疫情始于11月初，河南、河北、内蒙古一带较快蔓延，11月初以广州多区宣布解除疫情防控临时管控区为标志、广东主动迎接第一波冲击，12月初北京被动进入第一波冲击。目前上海、吉林、黑龙江等地疫情可能以较快的速度扩散。

● **风险提示：**国内疫情反复超预期，搜索指数有效性下降。

相关研究报告

《通胀连续下行，美联储内部可能分化——美国11月CPI点评—宏观经济点评》-2022.12.14

《杠杆仍加在企业端——11月社融数据点评—宏观经济点评》-2022.12.12

《后疫情时代来临，突出做好稳增长—宏观周报》-2022.12.11

目 录

| | |
|--------------------------------|---|
| 1、 防疫持续优化 | 3 |
| 2、 搜索指数映射疫情水平 | 3 |
| 2.1、 药品搜索指数最为有效 | 3 |
| 2.2、 区域搜索指数具有高频、领先、抗干扰特征 | 4 |
| 3、 利用搜索指数拟合疫情水平 | 6 |
| 4、 风险提示 | 7 |

图表目录

| | |
|------------------------------------|---|
| 图 1： 各地积极落实“新十条” | 3 |
| 图 2： 药品搜索指数较为有效 | 4 |
| 图 3： 新冠搜索指数有一定相关性 | 4 |
| 图 4： 发烧搜索指数从 11 月开始生效 | 4 |
| 图 5： “连花清瘟”搜索指数领先确诊病例 5-10 日 | 5 |
| 图 6： “新冠”搜索指数领先确诊病例 2-8 日 | 6 |
| 图 7： 连花清瘟指数与疫情节点较为吻合 | 6 |
| 图 8： 各地搜索指数与确诊病例拟合效果较好 | 7 |
| 图 9： 基于百度指数拟合日均确诊病例 | 7 |
| 表 1： 搜索指数与确诊病例相关性 | 5 |

1、防疫持续优化

11月以来我国防疫优化稳步迈进新阶段。11月11日“二十条”出台后，各地“因城施策”边际优化防疫政策。12月7日卫健委发布“新十条”，截至12月9日我们统计已有19个省市落实“新十条”。在“缩小核酸检测范围、减少频次”、“非必要不核酸”等导向下，以及多个省市取消常态化核酸，如何更好跟踪各地疫情水平以及判断第一轮冲击的达峰情况？

图1：各地积极落实“新十条”



数据来源：央视新闻、各地人民政府网、开源证券研究所

2、搜索指数映射疫情水平

表现疫情数据的传导链条为新冠→居民→卫健委，我们不妨针对前半段传导做文章。居民在感染或认为感染概率较高时，其可能发生的行为包括搜索新冠本体、新冠体征以及购买药品等。若能有效观测到前半段传导，理论上或可领先于卫健委的表现疫情数据。

2020年以来多篇学术论文基于谷歌趋势分析工具测算新冠疫情病例，例如2022年9月15日发表于《政策研究与战略管理》的一篇论文¹指出，“嗅觉丧失”、“味觉丧失”、“发烧”（按数量级计算）与官方病例显著相关。《中国-世卫组织新型冠状病毒肺炎（COVID-19）联合考察报告》对中国55924例实验室确诊病例的统计²显示，发烧、干咳、疲劳、咳痰、呼吸急促为前五大症状，出现比例分别为87.9%、67.7%、38.1%、33.4%、18.6%。因而“发烧”可能是居民搜索新冠体征的常用项。

2.1、药品搜索指数最为有效

陈沁博士在《各城市首轮感染高峰期预测》提出以发烧指数衡量疫情感染情况³，

¹ Lateef, B.A., Hossana, T., and Chinedu, W.O. (2022). Modeling COVID-19 incidence with Google Trends. Research Policy and Strategic Management.

² <http://www.nhc.gov.cn/xcs/yqfkdt/202002/87fd92510d094e4b9bad597608f5cc2c/files/fa3ab9461d0540c294b9982ac22af64d.pdf>

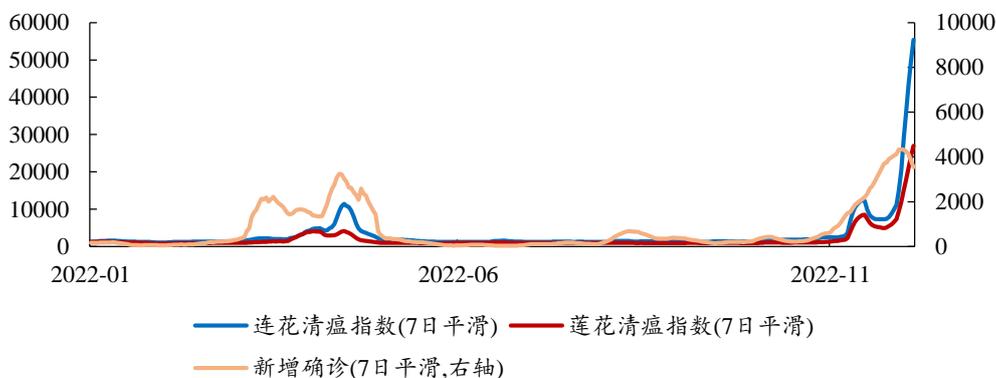
³ https://zhuanlan.zhihu.com/p/590989182?utm_medium=social&utm_oi=556272282197110784&utm_psn=1585332724964745217&utm_source=

请务必参阅正文后面的信息披露和法律声明

我们将其扩展为三类：搜索新冠本体、新冠体征以及购买药品，具体为“新冠”、“发烧”、“莲花清瘟”、“布洛芬”等关键词。

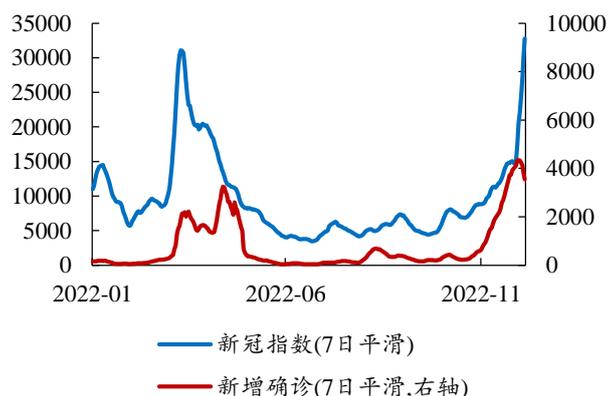
从与新增确诊的相关性来看，药品搜索指数>疫情搜索指数>体征搜索指数。有趣的是，“莲花清瘟”在三类关键词中相关性最好，附带的“莲花清瘟”同样具备映射疫情水平的作用。“二十条”发布后，莲花清瘟指数迅速上升而后下行，与疫情数据有小幅背离，在于防疫优化推进时药品包含了对未来疫情扩散的预期、因而提前上升；发烧指数在中国的有效性似乎并没有那么高，这可能与我国奥密克戎变异路径有关。2022年上半年主要是ba.2和ba.2.2变异株、下半年为BA.5和BF.7，前两者感染后临床表现相对温和、无症状感染者占比高，但BF.7感染症状包括高烧、持续性咳嗽、身体疼痛不适、头痛、喉咙痛等，与新冠奥密克戎早期亚型变异株感染后的症状类似。因此，高烧指数无法描绘前三个季度的疫情水平，但11月以来有效性显著提升。

图2：药品搜索指数较为有效



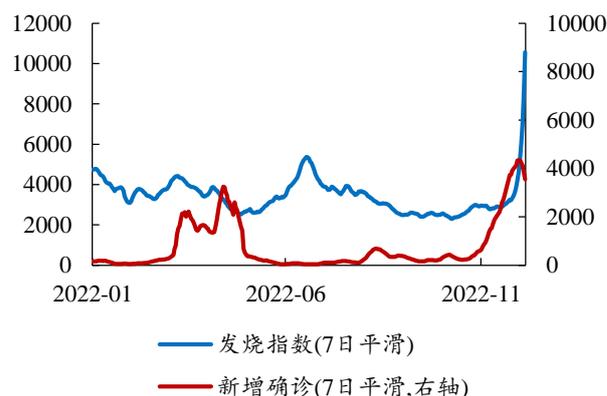
数据来源：百度指数、Wind、开源证券研究所

图3：新冠搜索指数有一定相关性



数据来源：百度指数、Wind、开源证券研究所

图4：发烧搜索指数从11月开始生效



数据来源：百度指数、Wind、开源证券研究所

2.2、区域搜索指数具有高频、领先、抗干扰特征

基于上述分析，“莲花清瘟”指数和“新冠”指数宜着重考察，我们发现具备

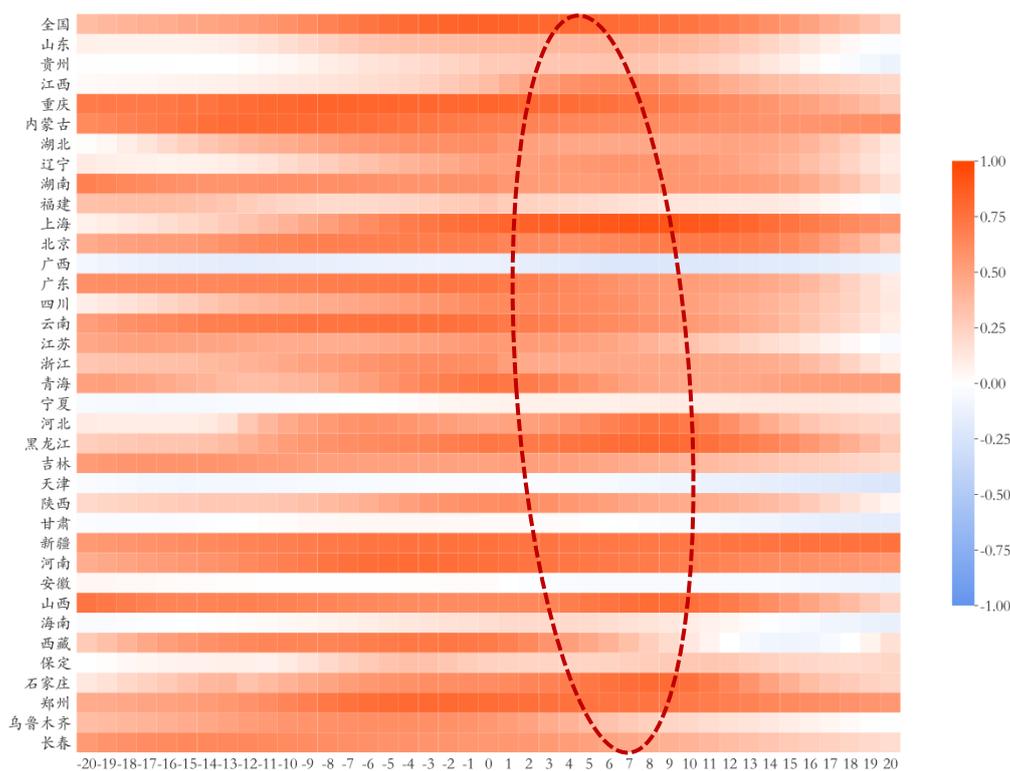
高频、领先性、抗干扰等特点。考虑到“新十条”对确诊病例的干扰，我们将时间窗口选择为2022年1月1日-2022年12月7日，并计算各搜索指数与确诊病例的相关性。从全国来看，连花清瘟指数、新冠指数、发烧指数与确诊病例的相关性分别为0.70、0.63、0.03；连花清瘟指数的领先期主要分布在5-10日、平均为7.1日，新冠指数的领先期主要分布在2-8日、平均为6.1日。

表1：搜索指数与确诊病例相关性

| 搜索指数 | 区域 | 相关性 | 相关性 (7日平滑) | 相关性 (最优领先期) | 相关性 (7日平滑, 最优领先期) |
|------|------|------|---------------|----------------|-------------------------|
| 连花清瘟 | 全国 | 0.70 | 0.72 | 0.84 | 0.84 |
| | 省市平均 | 0.42 | 0.50 | 0.51 | 0.57 |
| 新冠 | 全国 | 0.63 | 0.63 | 0.64 | 0.65 |
| | 省市平均 | 0.43 | 0.47 | 0.49 | 0.53 |
| 发烧 | 全国 | 0.03 | 0.03 | -0.06 | -0.02 |
| | 省市平均 | 0.08 | 0.18 | 0.10 | 0.21 |

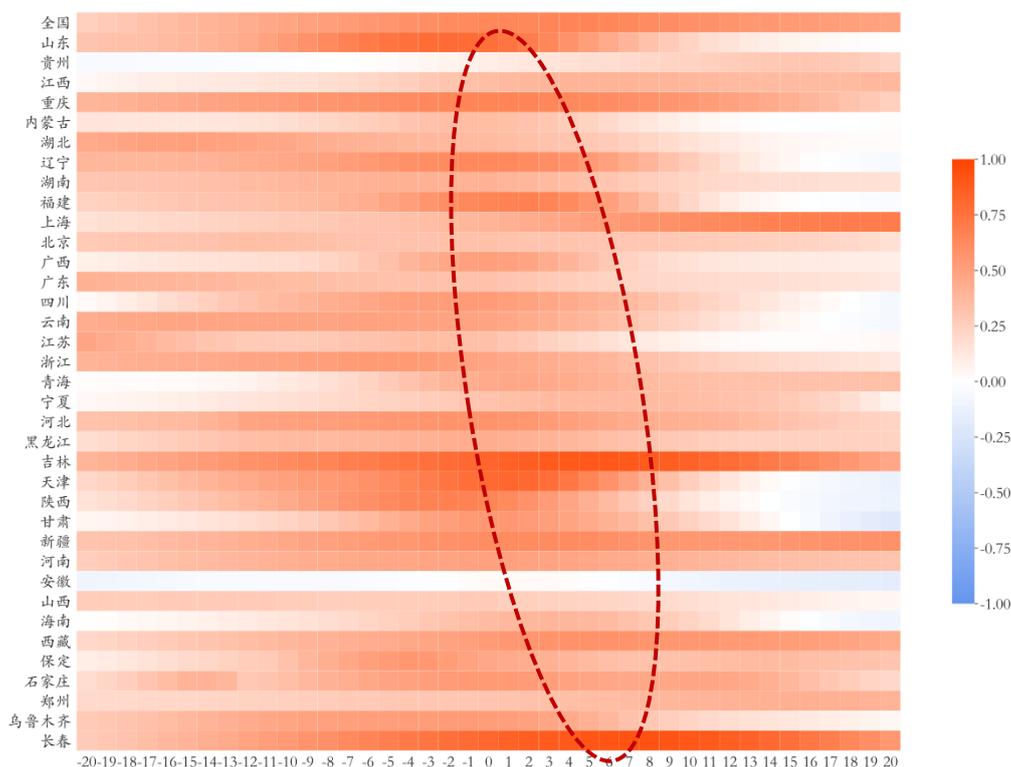
数据来源：百度指数、Wind、开源证券研究所

图5：“连花清瘟”搜索指数领先确诊病例5-10日



数据来源：百度指数、Wind、开源证券研究所；注：横坐标表示对搜索指数作平移处理，例如“1”列表示搜索指数前置1日时与确诊病例的相关性，“-1”列表示滞后1日，以此类推。

图6：“新冠”搜索指数领先确诊病例 2-8 日



数据来源：百度指数、Wind、开源证券研究所；注：横坐标表示对搜索指数作平移处理，例如“1”列表示搜索指数前置1日时与确诊病例的相关性，“-1”列表示滞后1日，以此类推。

表观疫情数据的收集、处理与披露受外部因素影响。当人力、医疗、设备资源匮乏时，表观疫情数据可能与真实疫情存在一定出入。从这个角度来看，搜索指数具有一定的抗干扰能力。

图7：连花清瘟指数与疫情节点较为吻合



预览已结束，完整报告链接和二维码如下：

https://www.yunbaogao.cn/report/index/report?reportId=1_50178



云报告
https://www.yunbaogao.cn

云报告
https://www.yunbaogao.cn

云报告
https://www.yunbaogao.cn