



北京大学

# 博士研究生学位论文

题目：应用于新冠疫情演化和脉搏波

分析的人体序参量理论研究

姓 名： 张磊

学 号： 1801111654

院 系： 工学院

专 业： 力学（生物力学与医学工程）

研究方向： 生物医学信息技术

导师姓名： 余振苏 教授

二〇二二 年 六 月



# **Study on human order parameters applied to COVID-19 evolution and pulse wave analysis**

Dissertation Submitted to  
**Peking University**

**In partial fulfillment of the requirements for the degree of  
Doctor of Engineering**

By  
**ZHANG Lei, M. Eng.**  
**(Mechanics(Biomechanics and Medical Engineering))**

Dissertation Supervisor: **SHE Zhensu, Professor**

Department of Mechanics and Engineering Science

College of Engineering

Peking University

June 2022, Beijing



## 版权声明

任何收存和保管本论文各种版本的单位和个人，未经本论文作者同意，不得将本论文转借他人，亦不得随意复制、抄录、拍照或以任何方式传播。否则，引起有碍作者著作权之问题，将可能承担法律责任。





## 摘要

人体是一个“一元二面多维多层次”的开放复杂巨系统，它不仅涉及到个体内部的多尺度耦合，同时与外部群体之间存在跨尺度的相互作用。面对这一复杂性的挑战，需要方法论的突破。

本文将朗道的“序参量”思想和余振苏等人发展的“湍流结构系综理论”的思想应用于人体复杂系统，旨在发展一套基于对称性分析的人体序参量理论，并在面向真实的二类人体复杂系统（即群体的新冠疫情演化和个体的脉搏波分析）中开展应用。

将上述理论应用于分析新冠疫情传播系统，得到远比传统传染病模型更为精确的 GHR 传染病动力学模型。本文发现，此次大规模抗疫过程中，除了传统传染病模型所刻画的由于病毒出现导致的对称破缺外，还存在国家干预所产生的对称破缺过程。本文提炼出国家三类干预过程中的三个关键序参量：隔离度 ( $G$ )，治愈度 ( $H$ ) 和抢救度 ( $R$ )，建立了 GHR 方程，得到疫情数据的精确支持。在将其应用于意大利和国内 24 个省份、不同流行波次、不同干预政策等真实疫情演化场景中，揭示了新冠疫情传播中的“序参量-转换速率-群体数量”三个层次相互耦合的演化规律。

进一步，本文根据 GHR 模型推导出粗死亡率满足 Logistic 的演化模式，并且在全球 52 个国家的第一波疫情的全周期死亡率数据中得到了验证。该模型的两个参数，将复杂的死亡率演化行为解耦为两个独立的过程，得到二个有实际应用价值的结论。一方面，本文发现各国死亡率演化早期的非存活者在 ICU 的持续时间 ( $\tau$ ) 与当地新冠疫情的持续时间之间具有普适的线性率关系，二者的比值大约为 7，这是此次新冠疫情展示的一个重要特征，如果正确，人们将能够根据疫情早期的 ICU 临床数据对疫情持续时间做出准确的预测。另一方面，死亡率演化晚期的饱和死亡率 ( $m_0$ ) 决定了死亡总人数，而本文发现它与国家中位数年龄、床位占有率为峰值之间分别具有幂律关系，定量解释了各国死亡率具有数量级差异的原因，并从中得出结论，若老龄国家（中位数年龄>38 岁）的病床数量增加一倍，可减少约 36% 的死亡人数。因此，本研究为下一波疫情准备医疗资源提供了定量的依据。

对于人体脉搏波系统，本文基于对称性原理猜测，人体脉搏波系统在长期演化中形成一个个性化的“常态”，同时在外内作用下不断发生新的对称破缺，形成了对常态的偏离，由此凝练出“偏离度”这个关键序参量来定量刻画个体健康态的变化。同时，本文根据脉搏波基本参数构建了“人体序参量相空间”，以之为基础对亚健康人群划分亚类和计算偏离度，构建了真正符合中医“知常达变”的脉诊原理的定量模型，突破了传统的大量统计平均方法无法开展个性化诊断的局限性。

进一步，本文将脉搏波序参量模型应用于人体健康指数的标定，结果显示它能够精确刻画个体常态的健康指数和个体日变的健康指数，且预测精度高于传统机器学习的结果。另外，这套模型框架被应用于刻画其他人体生理波动信号（如 GDV 信号），也显示了成功。这些结果表明，人体生理波动信号在序参量相空间中具有高度结构化的特征，这为开展融入医学经验和知识的机器学习提供了新的视角。

综上所述，本文基于“一元二面多维多层次”的人体复杂系统的本体论，发展了一套人体序参量理论。从群体的疫情演化到个体的脉搏波分析的应用结果显示，该理论能够定量刻画人体系统的自组织行为，为揭示主宰人体系统状态的数学原理，发现关键变量和普适常数，奠定了一定的基础。

关键词：复杂系统，人体系统，序参量，新冠疫情，脉搏波分析

## Abstract

The human body system is an open, complex and self-organized giant system. It involves not only multi-scale coupling within the individual but also cross-scale interactions with external groups. Facing the challenge of this complexity requires a methodological breakthrough. Based on the philosophical picture that human body owns a ‘single-centered, two-sided, multi-dimensional, multi-level’ structure, this study introduces the idea of order parameters which originated from Landau and developed by the Structural Ensemble Dynamics of turbulence to human complex systems, aiming to develop a symmetry-based theory of human order parameters. And this article applies it to two types of realistic human complex systems, the population evolution of COVID-19 and the individual analysis of pulse waves.

Applying the theory above to the evolution of COVID-19, a far more accurate epidemic dynamical model named GHR is obtained. This paper finds that there is another symmetry-breaking process generated by the national large-scale intervention of COVID-19, in addition to the symmetry-breaking due to the virus modeled by the SEIR model. We present three key order parameters of the national intervention process: governing force (G), healing force (H), and rescue force (R) to establish the GHR model, which is accurately supported by the epidemic data. The interaction between the three levels of ‘order parameters - conversion rate - population numbers’ in the spread of the epidemic is revealed by comparing the evolution of the epidemic in China and abroad in detail.

Further, a Logistic model of the crude mortality is derived from the GHR model and validated in the whole-cycle mortality data of the first wave across 52 countries. Two independent processes of the complex mortality evolution are decoupled by two model parameters, leading to two conclusions with practical applications. On the one hand, this paper finds a linear law between the ICU duration ( $\tau$ ) of non-survivors at the early stage and the local epidemic duration with a ratio of approximately 7, which is an important feature of COVID-19. And if correct, one would be able to make accurate predictions of the epidemic duration based on the clinical data from the ICU at the early stage. On the other hand, the saturation mortality rate ( $m_0$ ), which determined the total deaths at the late stage, obeyed two power-law with the median age and the bed occupancy respectively, quantitatively explaining the reasons for the magnitude differences in mortality across countries. And from this, the

conclusion was reached that about 36% of deaths could be avoided if the number of hospital beds in older countries (median age >38 years) were doubled. Therefore, the study provides a quantitative basis for preparing medical resources for the next wave of the epidemic.

For the human pulse wave system, this paper hypothesizes that the human pulse wave system forms an individualized ‘normal state’ during the long-term evolution based on the symmetry principle, while new symmetry breakdowns occur continuously under internal and external effects, resulting in deviations from the normal state. Thus, the key order parameter ‘deviation’ was developed to quantitatively describe changes in individual health status. At the same time, this paper constructs a ‘human order parameters phase space’ based on the basic pulse wave parameters, which are used to classify subcategories and calculate the ‘deviation’ for sub-healthy people. Therefore, a quantitative model is constructed which is in line with the principle of ‘knowing constant to achieve change’ in Traditional Chinese Medicine, breaking through the limitations of the traditional statistical averaging methods.

Further, the order parameters model of the pulse wave is applied to describe the individual's normal health index and daily health index, showing a more accurate description than results of the traditional machine learning method. In addition, this model framework was also successfully applied to describe other human physiological signals (e.g., GDV signal). These results show that human physiological signals are highly structured in the order parameters phase space, which provides a new perspective for developing the machine learning techniques incorporating medical experience and knowledge.

In summary, this thesis develops an order parameter theory of the human body based on the ‘single-centered, two-sided, multi-dimensional, multi-level’ philosophy of complex human system. The theory can quantitatively characterize the self-organizing behavior of human systems, laying a foundation for revealing the mathematical principles that govern the state of human systems and discovering the key variables and universal constants by applying the theory from population epidemic evolution to individual pulse wave analysis.

KEY WORDS: Complex Systems, Human Body System, Order Parameters, COVID-19, Pulse Wave Analysis

## 目录

<b>第一章 绪论</b> .....	1
1.1 复杂性科学研究概述 .....	2
1.1.1 欧洲学派 .....	2
1.1.2 美国学派 .....	3
1.1.3 中国学派 .....	4
1.1.4 挑战与机遇 .....	6
1.2 人体系统科学研究概述 .....	7
1.2.1 钱学森论人体科学 .....	7
1.2.2 人体系统科学概述 .....	7
1.3 本文的研究目标 .....	8
1.4 本文的组织结构 .....	9
<b>第二章 人体序参量理论研究</b> .....	11
2.1 朗道的序参量理论 .....	11
2.2 结构系综理论的突破 .....	12
2.3 人体序参量理论探索 .....	14
2.3.1 研究思路 .....	14
2.3.2 研究方法 .....	15
2.3.3 研究进展与特点 .....	16
2.4 应用于人体复杂系统 .....	17
2.4.1 群体：新冠疫情演化系统 .....	17
2.4.2 个体：人体脉搏波系统 .....	18
<b>第三章 应用于新冠疫情传播系统</b> .....	21
3.1 新冠疫情传播的复杂性 .....	21
3.1.1 疫情传播的建模研究 .....	21
3.1.2 已有研究的局限 .....	22
3.2 新冠疫情传播的多层次对称性刻画 .....	23
3.2.1 病毒出现后的对称性刻画 .....	23
3.2.2 国家干预后的对称性刻画 .....	25
3.3 新冠疫情传播的 GHR 模型 .....	26

---

3.3.1 序参量的解析表达 .....	26
3.3.2 GHR 模型 .....	27
3.3.3 模型求解及参数意义 .....	28
3.3.4 差分方程系统 .....	30
3.3.5 参数反演算法 .....	31
3.4 定量刻画国家干预力度 .....	33
3.4.1 疫情爆发中心地区的比较 .....	33
3.4.2 23 个省市的两类治愈度演化模式 .....	36
3.4.3 北京市在两波疫情中的成功经验 .....	39
3.4.4 定量模拟动态清零政策 .....	41
3.5 本章小结 .....	43
<b>第四章 应用于新冠死亡演化系统 .....</b>	<b>45</b>
4.1 新冠死亡率的复杂性 .....	45
4.1.1 三种死亡率及其影响因素 .....	45
4.1.2 关键问题：各国死亡率的巨大差异 .....	47
4.1.3 已有研究需要实现的突破 .....	47
4.2 新冠死亡率的序参量模型 .....	48
4.2.1 定性分类不同动力学特征 .....	49
4.2.2 定量刻画全周期演化规律 .....	50
4.2.3 模型参数的意义 .....	51
4.2.4 刻画 52 个国家的死亡率数据 .....	52
4.3 饱和死亡率的跨尺度特征 .....	54
4.3.1 与年龄、床位占有率的幂次律 .....	54
4.3.2 三类典型的演化模式 .....	56
4.3.3 国家干预措施的定量评估 .....	57
4.4 疫情持续时间的跨尺度特征 .....	58
4.4.1 ICU 时间 $\tau$ .....	58
4.4.2 无量纲系数 $k$ .....	59
4.4.3 普适的线性率 .....	60
4.4.4 预测疫情持续时间 .....	61
4.5 本章小结 .....	61
<b>第五章 应用于人体脉搏波系统 .....</b>	<b>65</b>
5.1 脉搏波研究的历史 .....	65

5.1.1 中医脉诊的研究 .....	65
5.1.2 现代脉搏波的研究 .....	67
5.2 脉搏波研究的挑战 .....	68
5.2.1 脉搏波研究的现状 .....	70
5.2.2 脉搏波研究的复杂性 .....	71
5.3 脉搏波系统的多层次对称性刻画 .....	72
5.3.1 个体常态的对称性刻画 .....	73
5.3.2 个体偏离态的对称性刻画 .....	74
5.3.3 知常达变的序参量解释 .....	75
5.4 提取多层次的脉搏波特征参数 .....	75
5.4.1 实验测量与数据采集 .....	75
5.4.2 数据预处理 .....	77
5.4.3 主导特征参数 .....	78
5.4.4 脉形特征参数 .....	82
5.5 脉搏波分析的序参量模型 .....	84
5.5.1 序参量相空间 .....	84
5.5.2 确定不同的亚类 .....	84
5.5.3 确定理想态和偏离度 .....	85
5.5.4 刻画人体健康数据 .....	85
5.5.5 模型的特点 .....	86
5.6 本章小结 .....	87
<b>第六章 应用于人体健康指数的标定 .....</b>	<b>89</b>
6.1 健康指数问卷 .....	89
6.2 刻画群体的稳定健康指数 .....	91
6.2.1 数据情况介绍 .....	91
6.2.2 建立序参量模型 .....	91
6.2.3 模型的验证 .....	91
6.2.4 与传统分类结果相比较 .....	93
6.2.5 各亚类的健康态特点 .....	94
6.3 刻画个体的日变健康指数 .....	96
6.3.1 数据情况介绍 .....	96
6.3.2 构建序参量模型 .....	96
6.3.3 普适的线性率 .....	97

6.3.4 各亚类的关系 .....	98
6.3.5 特征距离的时间演化 .....	100
6.3.6 日变健康指数的预测 .....	101
6.4 应用于分析其他人体波动信号 .....	102
6.4.1 数据情况介绍 .....	102
6.4.2 建立序参量模型 .....	103
6.4.3 普适的线性率 .....	104
6.4.4 不同亚类的特点 .....	105
6.5 本章小结 .....	106
<b>第七章 总结和展望 .....</b>	<b>109</b>
7.1 主要结论 .....	109
7.2 主要创新点 .....	110
7.3 研究计划 .....	110
<b>参考文献 .....</b>	<b>113</b>
<b>附录 A GDV 图像处理流程 .....</b>	<b>121</b>
<b>附录 B 脉搏波与健康指数的原始数据 .....</b>	<b>127</b>
<b>个人简历、在学期间发表的学术论文与研究成果 .....</b>	<b>139</b>
<b>致谢 .....</b>	<b>141</b>
<b>北京大学学位论文原创性声明和使用授权说明 .....</b>	<b>147</b>
<b>学位论文答辩委员会名单 .....</b>	<b>149</b>
<b>博士学位论文答辩委员会决议书 .....</b>	<b>151</b>
<b>提交终版学位论文承诺书 .....</b>	<b>153</b>

## 第一章 绪论

我们生活在这样一个奇妙的世界。苏东坡先生说，“盖将自其变者而观之，则天地曾不能以一瞬；自其不变者而观之，则物与我皆无尽也”。“变”和“不变”，这是一个永恒的哲学话题，同时也是一个重要的科学问题。

20世纪前的四百多年里，现代科学创造了巨大成就。这些成就建立在一种称之为“还原论”的方法之上，即针对研究对象，“用经验分析的方法，把整体分解为部分，把高层次还原到低层次”<sup>[1]</sup>来认识事物。这一方法在物理、化学等简单系统的研究中十分奏效，以至于让人们觉得，宇宙一切事物的运动，都已经被决定了。人们曾假想通过牛顿定律就能知晓宇宙过去与未来的“拉普拉斯妖”，就是一个很好的例证。然而，就在一片欢欣鼓舞中，自然科学却孕育出了一场至今仍然在发生过程中的科学革命。

20世纪初，相对论和量子力学的建立，极大地冲击了原来人们的绝对时空观和还原方法论。在微观层次上，“海森堡测不准原理”宣布了主客观分离和确定性的失效<sup>[2]</sup>；在宏观层次上，混沌现象让人们认识到，即使对于十分确定的简单系统，初始条件的微小改变也会导致它发生不可预测的演化<sup>[3,4]</sup>。人们几百年赖以坚信的“简单性范式”的基石开始动摇了，一些有识之士的反思开始走进人们的视野。例如，量子力学奠基人之一的薛定谔，在他著名的系列演讲<sup>[5]</sup>中反思了传统科学图景的严重缺陷，“这个图景使你把世界的整个表现都作为机械时钟的行为，……我们已经使用了极其简化的手段，把我们的人性排除了，移走了”。

人们把目光转向了真实世界的系统，发现实际上我们生活在一个充满复杂性的世界。从植物的光合作用，到动物的集体行为，从江河湖海的湍流现象，到天气气候的多尺度运动，这些过程中都涉及十分复杂的动力学过程，却总是能够在系统层次上展示出令人印象深刻的有序性。

“复杂性”这一概念被提了出来，人们不得不去正视各个研究领域中的“复杂性问题”。复杂性研究的先驱贝塔朗菲曾说道，“我们被迫在一切知识领域中运用‘整体’或‘系统’概念来处理复杂性问题”<sup>[6]</sup>。由此可见，复杂性研究也是整个科学系统自身发展演化的新阶段<sup>[7]</sup>。

直面复杂性的挑战，一般系统论(General System Theory)<sup>[6]</sup>、耗散结构论(Dissipative Structure Theory)<sup>[8]</sup>、协同学(Synergetics)<sup>[9]</sup>、复杂适应系统(Complex Adaptive System, CAS)<sup>[10]</sup>、开放的复杂巨系统(Open Complex Giant System, OCGS)<sup>[11]</sup>等理论被相继提出。一场关于“复杂性”研究的盛宴徐徐展开。为了更好地论述本文的研究主题，本章将在简要回顾复杂性研究代表成果的基础上，提出对人体复杂系统研究的价值和意义。

## 1.1 复杂性科学的研究概述

在过去不到一百年的时间里，复杂性科学经历了不同阶段的发展浪潮，形成了比较有影响力的三大典型学派<sup>[7]</sup>。由于篇幅所限，我们只能选取一些与本文主题比较相关的代表性工作进行概述，以勾勒出一个简要的发展脉络。

### 1.1.1 欧洲学派

欧洲学派的主要贡献之一是提出了“自组织理论”(Self-organizing Theory)。这一理论主要包括耗散结构理论和协同学等，而其开端要追溯到贝塔朗菲(L.Von Bertalanfy)的工作。贝塔朗菲是一位奥地利的生物学家，他在生物研究中感到将生物有机体不断细分的还原论方法会越来越模糊对生物整体的认识——“流行的机械论方法所忽视的并起劲地加以否定的，正是生命现象中最基本的那些东西”<sup>[6]</sup>。因此，他强调要“把有机体作为一个整体或系统来考虑”<sup>[6]</sup>，而“生物学的首要任务必然是发现生物系统（在有机体的各个层次上）的规律……一种生物体的有机理论或系统理论”<sup>[12]</sup>。在这之后，他把这些思想推广到更一般的系统，提出了从整体上研究问题的整体论方法，建立了“一般系统论”。虽然还没有很好的整体论的方法体系，但是“一般系统论”较早地开展了对还原论方法的质疑，并建立起“涌现”“组织”“开放系统”等至今仍然在复杂系统研究中使用的重要概念。

贝塔朗菲的一般系统论是在 20 世纪 40 年代提出来的，其理论背景更受到平衡态的或近平衡态的热力学所影响，因此“过分强调了整体性、有序性和统一性的观念，而完全否定了局部性、无序性和分散性的观念”<sup>[13]</sup>。而普利高津(Ilya Prigogine)则是首先将注意力从平衡态转移到远离平衡态的非线性区，提出了“有序来自于无序”的口号，建立了“耗散结构理论”。

普利高津是一位物理化学家和理论物理学家，创建了“布鲁塞尔学派”。早期的热力学理论描绘了宇宙热寂的演化结果，人们更关注于对平衡态的研究，认为无序的非平衡状态是不能产生出有序结构的。而普利高津采用了动力学的眼光，认为对自然界常见的“不可逆过程”的研究才是更为重要的。根据昂萨格倒易关系，普利高津证明了系统在近平衡的线性区满足最小熵产生原理，不可能形成稳定的有序结构。进一步，拓展到远离平衡态的非线性区，开放系统在外界“负熵流”的作用下，可以通过自组织达到一种有序状态。这种时间、空间或功能有序的稳定结构，必须要有外界持续的负熵输入，因此被称之为“耗散结构”。而要形成耗散结构，必须是与外界有稳定的物质、能量和信息交换的开放系统，系统处于远离平衡态的非线性区，系统本身存在涨落，以及且系统内部子系统之间要存在非线性相互作用<sup>[14]</sup>。

耗散结构理论回答了系统在外界作用下是如何从无序走向有序的问题，展示了不同于热力学第二定律所规定的演化模式，给人们带来了新的希望。普利高津也因此获

得了 1977 年的诺贝尔化学奖。

但是，真正尝试对系统的自组织给出定量数学描述的，是协同学的创始人赫尔曼·哈肯（Hermann Haken）。协同学的研究对象也是远离平衡态的开放系统。与普利高津所不同，哈肯认为从有序到无序的转化不在于“非平衡态”，而是在于系统的各个子系统之间的协同作用。

那么，如何描述这种协同作用呢？哈肯认为，“（熵）作为处理自组织结构的工具是太粗糙了……需要有其他的途径”<sup>[9]</sup>。他创造性地将朗道的“序参量”思想引入到了复杂系统研究之中，从而突破了对“协同作用”进行数学描述的难题。这一惊人的创造，实际上与他的理论物理学的背景密切相关。哈肯曾长期深入研究激光系统，他将激光产生的过程与超导平衡相变的过程进行类比，发现二者存在着深刻的相似性。因此，哈肯认为非平衡系统中的有序结构的生成过程，是平衡系统中相变过程的拓展。所以很自然地，引入了平衡相变理论中的序参量的概念和绝热消去原理，来刻画非平衡系统的演化行为<sup>[15]</sup>。

具体来说，系统演化方程中随时间变化快的状态变量，受到慢变量的主导（协同学中称为“役使原理”）。进一步，通过扰动等数学方法消去快变量，留下的慢变量就是主导系统演化的序参量。通过研究序参量的演化方程，就可以分析系统的各种非平衡定态和不稳定性。通常使用的系统演化方程有主方程、郎之万方程、福克-普朗克方程和广义的金兹堡-朗道方程等。

协同学使得自组织理论可以定量地研究系统的演化行为，并且还被广泛地应用于经济<sup>[16]</sup>、管理<sup>[17,18]</sup>和社会<sup>[19]</sup>等各种复杂系统中。这一方法的局限性在于，如果宏观序参量过多或无法建立有效的序参量方程，协同学也只能提供定性分析<sup>[7]</sup>。更重要的是，协同学的分析很大程度上受限于所使用的演化方程。但目前采用的这些演化方程，对于真实世界的复杂系统而言过于简化，因此距离复杂系统的真实演化情况较远<sup>[20]</sup>。

### 1.1.2 美国学派

1984 年，三位诺贝尔奖获得者：盖尔曼（Murray Gell-Mann）、安德森（Philip Anderson）和阿罗（Arrow）联合一批不同领域的学者，建立了第一个专门以研究复杂性为主题的研究机构，即著名的圣菲研究所（Santa Fe Institute, SFI），并逐渐成为了著名的复杂系统研究中心之一。

SFI 的主要研究对象是复杂适应系统。复杂适应系统理论是美国的霍兰德（John H. Holland）教授在 1994 年时正式提出的。这一理论的核心概念是“具有适应能力的主体”（Adaptive Agent）。它早已不是之前系统研究中所使用的元素、子系统等概念的范畴，而是具有适应性、学习性的“活”的个体。而复杂适应系统就是由大量的这样的“主体”所构成的系统。其中每一个“主体”所处的主要环境，就是由其他“主体”所组成环境。

“主体”在与环境的相互作用中，不断地学习并做出适应性的改变，从而可能在更大尺度上“涌现”出系统层次的复杂演化行为。因此，霍兰德提出了“适应性造就复杂性”的观点<sup>[10]</sup>。

复杂适应系统的建模方法主要是基于主体的模型（Agent-based model）。SFI 开发了一个公共的模拟平台——Swarm，来支持研究者对复杂适应系统开展建模和模拟工作。苗东升教授<sup>[21]</sup>曾总结，SFI 对复杂性研究的主要贡献是科学地证明了“复杂性”是通过系统的自组织而涌现出来的，从而把“涌现”的概念从哲学认识走向了科学，使其成为系统科学和复杂系统研究的基本概念。

其次，SFI 的另一项研究主题是“人工生命”（Artificial life）。这一概念最初是由克里斯托弗·朗顿（Christopher Langton）于 1986 年提出，旨在通过计算机设置的人工环境来模拟生命系统的进化。这一研究推动了元胞自动机的发展，获得了许多著名的探索人工生命演化的模型，如鸟群模型、蚁群模型等。

目前，基于主体的模型仍然是国际复杂系统研究的主流<sup>[22]</sup>。但是，与元胞自动机等二层次的模拟方法一样，这种方法面临着主体简单、层次不够丰富，难以获得自上而下的指导等挑战<sup>[23]</sup>。正如学者指出<sup>[13]</sup>，SFI 试图建立一种新型的还原论，其还原方式不是从整体还原到个体，而是还原到个体相互作用的简单规则。

最后，“复杂网络”的研究掀起了复杂系统研究的新一轮浪潮。1998 年，Watts 和 Strogatz 在《Nature》上发表了题为“‘小世界’网络的集体动力学”的文章<sup>[24]</sup>，从数学上首次定义了“小世界网络”，揭示了复杂网络的小世界特征，也被称为 W-S 模型。第二年，Barabási 和 Albert 在《Science》上发表了“随机网络标度的涌现”的文章<sup>[25]</sup>，提出了复杂网络的无标度性质，也被称为 B-A 模型。这两篇里程碑式的论文，开启了复杂网络研究新纪元。目前，复杂网络研究发展迅速，被广泛地应用于计算机、电力网络、气候科学和流行病学等领域<sup>[26]</sup>。

### 1.1.3 中国学派

中国传统文化中蕴含着丰富的系统观念。特别是中华传统医学作为其中的杰出代表，被称之为“第一门复杂性科学”<sup>[27-29]</sup>。

前文提到的复杂性科学的研究的代表学者们，也都对此做出了中肯的评价。例如，普利高津认为，“中国道家对人类、社会和自然之间有着深刻的理解，这对西方哲学家和科学家始终是个启迪的源泉”<sup>[8]</sup>。哈肯提出，“协同学含有中国基本思维的一些特点。事实上，对自然的整体理解是中国哲学的一个核心部分”<sup>[9]</sup>。霍兰德则说，“真正综合两种传统——欧美科学的逻辑—数学方法与中国传统的隐喻类比相结合——可能有效地打破现存的两种传统截然分离的种种限制。在人类历史上，我们正面临着复杂问题的研究，综合两种传统或许能使我们做得更好”<sup>[10]</sup>。

但是，综合集成中西方传统，这本身就是一件充满复杂性挑战的事情。钱学森先生在这方面做出了开创性的工作，在他的带领和推动下，经过几十年的努力，开创了复杂性研究的中国学派。

20世纪80年代初——大约与圣菲研究所成立的同一时间，钱老开始了近三十年的从系统思想到系统实践的创新过程。特别需要说明的是，这个过程本身也是用系统学的方法来开展的。这就是钱老创建的“系统学讨论班”的形式。通过这种形式，来综合集成各方面专家的意见，提炼新的思想。从1986年到1992年，钱老参加了所有讨论班的学术活动<sup>[30]</sup>。在此过程中，钱老还与全国各地、各行各业的专家学者们开展了密集的书信交流。据目前已出版的不完全统计，这一时期的书信至少是在5000封以上<sup>[31]</sup>。由此对钱老在这一时期的科学探索之高远，学术创造力之旺盛，以及系统学创建本身之复杂等等，可见一斑。参考前人的总结<sup>[30,32]</sup>，我们从以下几个方面，简要地总结了钱老等人所创建的“开放的复杂巨系统理论”的主要内容：

在系统认识论方面，提出了系统的一般分类和系统学的四个层次。系统可以分为简单系统、简单巨系统、复杂巨系统和特殊复杂巨系统<sup>[33]</sup>。而对于复杂系统，又都需要从工程实践、到技术科学、基础科学，最后上升至哲学——这四个层次上开展研究。由此，钱老构建了包含十一个门类的现代科学技术体系。其中，自然科学、社会科学和数学科学都已经有了较良好的基础，而其他门类还尚未完成，却都是社会发展的必然需求。

在系统方法论方面，提出了“从定性到定量的综合集成法”。这一方法论“本质上是科学和经验的结合”<sup>[11]</sup>，被认为要起到“一切新学科的微积分”<sup>[34]</sup>的作用。具体来说，通过“人机结合，以人为主”的思维方式和研究方式，构建高度智能化的人机结合的融合体系，综合集成专家的意见形成假设，开展多参数的定量建模工作，并且需要在实践中不断来回检验和迭代。

在系统实践论方面，提出了“专家研讨厅体系”和“总体设计部”的系统工程的实践形式。很自然地，面对实际要解决的复杂问题，要运用上述综合集成方法，就需要有多方面的专家来进行充分地研讨和设计。这不仅仅是参与决策和领导，而是要贯彻于整个系统工程实践中的整个过程中。

最后，要完成上述理论的构建，并付诸于实践中的复杂问题的解决，其主体就需要是全面发展的人才，因此提出了大成智慧、大成智慧工程、大成智慧教育等新概念。希望在具体的复杂系统科学的研究和工作实践中，通过大成智慧教育，培养出“帅才型”的大成智慧者，在社会主义建设中发挥至关重要的作用，从而也是某种对“钱学森之间”的回应。

综上所述，钱学森开创了复杂巨系统的科学与技术这个新领域，并建立了一套完整 的认识论、方法论和实践论框架。在他的带领和推动下，一批学者在复杂系统的哲学

[<sup>21</sup>]、复杂系统的统计力学<sup>[35]</sup>、知识系统工程<sup>[36]</sup>、开放的复杂智能系统<sup>[37]</sup>、专家挖掘等综合集成技术<sup>[38]</sup>和地理系统科学等方面开展了深入的研究，取得了重要的进展。在这里，特别要介绍马蔼乃教授和余振苏教授在继承和发展钱老系统科学思想方面所做的工作。

马蔼乃教授在长期从事地理科学研究的过程中，逐渐认识到必须把地理系统作为一个复杂系统来研究，并且必须要有定量的数理模型来支撑。她将钱老的系统科学思想应用于地理学中，构建了包含地理哲学<sup>[39]</sup>、地理基础科学<sup>[40]</sup>、天地人机信息一体化网络系统的地理信息科学<sup>[41]</sup>和地理系统工程<sup>[42,43]</sup>一共四个层次的地理系统科学，填补了现代科学技术体系中“地理科学”这一大类的空白<sup>[44]</sup>。并且，相对于钱老倡导发展的其他科学部门而言，“地理科学”是发展最为完善的一个科学门类，这对于推动复杂性地理科学的发展具有重要的意义。

余振苏教授提出，钱老的系统科学思想的深入发展，需要复杂系统的本体论的支撑，并构建了一套“一元二面多维多层次”的复杂系统学新框架<sup>[32]</sup>。具体来说，首先在本体论方面，提出以“一元二面多维多层次”的具体特征来充实开放的复杂巨系统的理论框架。在认识论方面，提出了“认识主客体的相对复杂性原理”“认识真理性的时空相对性原理”“理性知识的层次性原理”和“理性知识的知识宝塔原理”等复杂系统的四条认识论原理。在方法论方面，提出了“知识宝塔方法”，作为“从定性到定量的综合集成法”的具体实施方法。进一步，余振苏教授将这套研究框架，应用于对人体复杂系统和社会复杂系统的研究，建立了相应的研究框架，初步展示了该套系统学新框架的应用价值。

#### 1.1.4 挑战与机遇

从上面简要的介绍中可以看出，复杂系统研究的三大学派各自具有鲜明的研究特点，显示出复杂性研究本身的勃勃生机。但同时，目前复杂性研究面临着以下三个方面的挑战：

第一，研究问题的复杂度不足。例如，目前许多复杂性研究的对象实际是人工模拟的复杂系统。虽然较一般的物理系统的复杂性较高，但是与真实的复杂系统相比，其复杂性已经大大降低。所以，这对于研究简单规则下的系统发生“涌现”等演化行为很有帮助，但是距离真实复杂系统的应用还有较远距离。

第二，研究方法的还原论偏向大。例如，目前流行的基于多主体的建模工作，实际上仍然是一种自下而上的模拟，缺乏上层目标的设计和反馈。这也是与真实复杂系统相差最大的地方。

第三，缺乏统一且定量的研究框架。中国学派虽然提出了较为完善的、统一的复杂系统研究框架，但是在具体内涵上还缺乏定量的模型支撑。例如，如何从定性走向定

量？如何进行综合集成？这些都是在面对具体的复杂系统研究时，需要回答的重要问题。

但同时，我们也看到，随着人工智能、大数据等信息科技的发展，复杂网络等研究掀起了新一轮的研究热潮<sup>[45]</sup>，复杂性研究已经超出了其自身的范畴，而是更多地参与到传统学科复杂问题的跨学科研究中去<sup>[46]</sup>。在这个集中合作的研究过程中，笔者认为，必将总结出对于复杂系统研究的新范式，一个复杂性研究的成果爆发期即将到来。

## 1.2 人体系统科学的研究概述

面对上述复杂性研究的挑战，我们采取的研究策略是，针对某一类真实的复杂系统，对其丰富的演化行为开展定量刻画，并由此总结出普适的研究范式。

一方面，由于我们当前比历史上任何时候都具有更丰富的关于“人”的真实演化的大数据。另一方面，不同个体的活动又构成了群体的复杂系统演化。这样，就形成了一个从个体到群体的多层次研究框架。所以，我们选择了“人体复杂系统”作为研究对象，以此来发展复杂系统研究的新范式。在此之前，本节首先简要概述人体复杂系统科学的研究进展。

### 1.2.1 钱学森论人体科学

钱学森先生从开放的复杂巨系统理论出发，提出了“人体科学”的概念：“人体是一个开放的巨系统，他的特征是人体的功能状态，包括一些特殊的人体功能状态。人体科学就是研究人和人在客观环境中所处功能的学问。”<sup>[47]</sup>

进一步，钱学森提出，人体科学作为“现代科学技术体系”中的一个部门，应该包括“人天观哲学”“基础科学”“技术科学”和“工程技术”四大层次的内容建设。在这个框架下，来综合集成古今中外、自然科学和社会科学、理想的和经验的有关“人体”的所有知识。

这一学科的建立，将突破机械的科学观和还原论的思维观的束缚，对人体这一复杂系统开展深入系统的研究。钱学森指出，这“可能导致一场 21 世纪的新的科学革命，也许是比 20 世纪初的量子力学、相对论更伟大的科学革命”，将“使我们认识客观世界和改造客观世界来一次更大的总的飞跃，这难道不是第二次文艺复兴吗？”<sup>[47]</sup>

### 1.2.2 人体系统科学概述

余振苏教授在以往对湍流、生物进化、中医系统论、运动员训练等复杂系统研究和实践的基础上，继承并发展了钱学森的人体科学思想，开展了人体系统科学框架的构建<sup>[23,32]</sup>。

首先，他提出了人体系统是一个“一元二面多维多层次”的本体论新模型。所谓“一元”，可以认为是人体系统的自组织中心，是最不容易把握的。但是，在本体论的指导下，可以通过多个维度的“二面”的综合集成来开展认识。

例如，余振苏提出，可以通过“虚实”“动静”和“内外”这三对“二面”来对人体系统的自组织中心，进行初步的刻画：

虚实二面：“实”指的是人体的生理结构，而“虚”是人的意识活动。钱学森认为，人体不同于一般生物体的重要特点在于，人的“意识又能作用于人体本身，是‘意识反馈’”。这也就是人体科学区别于一般生命科学的特征。因此人体科学的研究要把握物质与精神、客观与主观、大脑与意识的辩证关系”<sup>[48]</sup>。在人体自组织中心的调配下，人体系统的“虚”与“实”实现了和谐的互动。

动静二面：“动”指的是人体系统无时无刻不处于多个尺度的运动之中，如神经的放电、蛋白质的折叠、细胞的分裂、血液的流动等，但是在自组织中心的作用下，这些多尺度的运动能够耦合起来，形成一种更高级的有序态，这就是人体系统“静”的一面。以往的研究更多关注在不同尺度下的运动，而忽视了“静”这一面的有序态才是更加直接地体现了人体自组织中心的作用。

内外二面：“内”指的是人体系统本身，“外”则是人体系统所处的外环境。人体系统与外环境之间时刻处于物质、能量和信息的交换之中。这种高度开放性能够一直保持着一种有序状态，就是人体自组织中心的作用结果。

类似于上述的多个二面，就构成了对人体系统的多维度刻画。在不同尺度下，人体系统还具有典型的多层次结构。理解人体系统的多层次性，开展人体系统跨层次耦合的研究，是人体系统研究的重点。

其次，余振苏还提出了研究人体系统科学的方法论——“多层表述，逐级定量，多次迭代，逐步近似”，进一步将“从定性到定量的综合集成法”进行了更加具体化和可操作化的发展。

最后，余振苏还提出了人体系统的自组织原理、开放性原理、层次结构原理、能量原理和进化原理等五大原理，并将其应用于开发人体系统科学的应用技术和指导人体系统科学的工程实践<sup>[49-51]</sup>。

综上，余振苏等完成了人体系统科学框架的构建。这一理论的构建，发展了钱老提出的现代科学技术体系中“人体科学”这一大类的四个层次的具体内容，为开展人体复杂系统的科学的研究和工程实践奠定了坚实的基础。

### 1.3 本文的研究目标

综上所述，钱学森复杂系统思想和人体系统科学框架，为人体复杂巨系统的研究

提供了重要的方法论的指导。但是其中的具体内涵和普适性原理，还需要从对具体的复杂系统研究实践中，进行总结提炼。正如钱学森先生所说，“要建立开放复杂巨系统的一般理论，必须从一个一个具体的开放复杂巨系统入手。……只有从一个一个具体的开放复杂巨系统入手进行研究，当这些具体的开放复杂巨系统的研究成果多了，才能从中提炼出一般的开放复杂巨系统理论，形成开放的复杂巨系统学，作为系统学的一部分”<sup>[11]</sup>。

特别地，随着移动互联网、大数据、人工智能等新一代信息技术的快速发展，人们比以往任何时候都积累了更多的关于各种开放复杂巨系统的客观数据和处理技术。因此，针对具体的“开放的复杂巨系”开展深入的研究，提炼出背后的一般方法和科学原理，是正当其时，更是时代所需。

因此，本文的研究目标，就是要针对具体的人体复杂系统，运用上述人体系统科学的方法论，在解决某些重大问题的过程中，探索人体复杂系统研究的具体方法。这就是本文所要研究和建立的“人体序参量理论”，详细内容请见本文的第二章。

具体来讲，本文选取了人类文明史上所罕见的“新冠疫情演化系统”，以及与人体健康密切相关的“人体脉搏波系统”这两大系统，作为具体研究的切入点。从新冠疫情传播的定量刻画，到各国死亡率巨大差异的定量解释，再到人体生理波动信号的建模理解，这些都是目前全人类所关注的重大课题，具有重要的现实意义。同时，这些重大课题已经在前期研究过程中积累了丰富的数据和经验，也为本研究提供了重要的研究基础。

更重要的是，这两大系统正好构成了一幅从群体数量演化，到个体健康变化，再到人体内部脉动量涨落这样一个多层次的人体复杂系统研究图景。这对于探索人体复杂系统研究的一般方法，具有更为普适的重要意义。

## 1.4 本文的组织结构

本文的主要研究内容是针对“新冠疫情演化”和“人体脉搏波”这两大具体的开放复杂巨系统，构建人体序参量理论，探索人体复杂系统研究的具体方法。

第一章是绪论，首先回顾了复杂性科学研究历史中的代表成果，特别是人体复杂系统研究的主要内容和最新进展，由此提出了本文的研究目标和内容

第二章是本文的理论基础，通过继承朗道的序参量思想，特别是根据湍流研究突破带来的启示，本文尝试构建了人体序参量理论，并初步总结了其内容和特点，并且介绍了如何应用于具体的人体复杂系统研究。这为后面章节分别开展具体的建模研究奠定了基础。

第三章是围绕“新冠疫情传播系统”开展的工作。通过分析新冠疫情传播中的多次

对此破缺过程，本文提炼了隔离度（G）、治愈度（H）和抢救度（R）这三个影响疫情演化的关键序参量，并建立了新的传染病动力学模型。并且应用于不同特征地区、不同流行波次、不同干预政策等真实疫情演化场景的刻画和分析中，发现了新冠疫情演化中存在的“序参量-转换速率-种群数量”三个层次相互耦合的自组织规律。

第四章则是在第三章基础上，专门讨论了“新冠死亡率演化系统”。本文通过建立序参量模型，成功刻画了在第一波新冠疫情中的 52 个国家的全周期死亡率的演化行为。并且通过建立模型序参量与医疗和社会序参量之间的定量关系，发现了其中存在的空间和时间的跨尺度规律。特别是，本文发现了疫情早期的非存活者的 ICU 持续时间和疫情整体的持续时间之间存在普适的线性率，这为疫情早期开展准确预测，提供了重要支撑。

第五章是针对“人体脉搏波系统”开展了序参量建模工作。本文基于对称性原理猜测，人体脉搏波系统在长期演化中形成一个个性化的“常态”，同时在外作用下不断发生新的对称破缺，形成了对常态的偏离，由此凝练出“偏离度”这个关键序参量来定量刻画个体健康态的变化。同时，本文根据脉搏波基本参数构建了“人体序参量相空间”，以之为基础对亚健康人群划分亚类和计算偏离度，构建了真正符合中医“知常达变”的脉诊原理的定量模型，突破了传统的大量统计平均方法无法开展个性化诊断的局限性。

第六章是将脉搏波序参量模型应用于人体健康指数的标定。结果显示它能够精确刻画个体常态的健康指数和个体日变的健康指数，且预测精度高于传统机器学习的结果。另外，这套模型框架被应用于刻画其他人体生理波动信号（如 GDV 信号），也显示了成功。这些结果表明，人体生理波动信号在序参量相空间中具有高度结构化的特征，这为开展融入医学经验和知识的机器学习提供了新的视角。

第七章是总结与展望，总结了全篇工作的结论和创新，对进一步的研究工作进行了计划和展望。

## 第二章 人体序参量理论研究

20世纪物理学的几个突破性进展，都与“对称性”密切相关。基于“对称性”，爱因斯坦创立了相对论，杨振宁创立了现代规范场论。而朗道通过提出“序参量”的思想，建立了基于对称性分析构建序参量方程的新范式，在相变、超导、超流等方面的研究都获得了诺贝尔奖级别的突破。

余振苏等继承了朗道的思想，将“序参量”这一概念拓展到“序函数”，通过对称性分析，建立了湍流边界层的“结构系综理论”，实现了“湍流百年难题”的突破。本文正是在这一方法论的指导下，开展了针对人体复杂系统的序参量研究。

### 2.1 朗道的序参量理论

1937年，著名物理学家朗道（Lev Davidovich Landau）在研究相变问题时发现，“液体与晶体之间本质上是不同的，液体是各向同性的，而晶体相反，是各向异性的。从晶体到液体，或一种晶体到另一种不同对称性的晶体之间的每一次相变，是和某种对称元的消失或出现密切相关的。”<sup>[52,53]</sup>由此，朗道提出，由于对称性的变化导致了相变，而表征相变性质的决定性因素是“对称元”。为此，朗道提出了一个新的概念——“序参量（Order Parameter）”——来表征系统的宏观有序程度，并在此基础上建立了“二级相变理论”。

朗道理论是一个“平均场”的理论，微观粒子之间复杂相互作用带来的涨落被平均后，剩余的宏观的有序部分，就可以被“序参量”来表征。例如，铁磁体在高于居里点温度时，热运动打破了原子的有序状态而变得无序，从而在宏观上表现出磁性的消失，自发磁化强度为零。而当低温时，原子之间的相互作用又使系统能量降低而处于一种集体有序排列的状态，并在宏观上表现出磁性的恢复，此时的自发磁化强度为一有限值。从而，通过“自发磁化强度”这个量，我们就可以判断系统的有序程度——这个量就是“序参量”。

1950年，朗道和金兹堡基于序参量的思想，建立了超导的唯象理论。他们认为，在临界温度处由正常态向超导态的转变是一个二级相变，根据对称性分析，确定了序参量是超导电子的有效波函数 $\psi$ 。根据相变理论，将自由能展开为序参量的幂级数，并根据自由能取最小值的要求，推导出了著名的“金兹堡-朗道方程”（Ginzburg-Landau theory）。

这一理论被誉为“物理直觉的惊人绝技”<sup>[54]</sup>。因为当时人们对于微观的过程并不是很清楚，但是金兹堡-朗道仅通过对称性的分析，就确定出超导序参量一定是一个与

宏观对称破缺相联系的复数，从而不依赖任何微观理论就写出了正确的方程。金兹堡-朗道理论直接导致了第二类超导体的发现，还成功解释了超流性、位移型相变、液晶中的相变等。虽然后来 BCS 超导微观理论成为人们理解超导现象的基础，但是由于它只适用于超导材料为空间均匀的情况，所以更为一般适用的金兹堡-朗道理论仍然是人们理解超导现象的重要途径，乃至被誉为“鉴于技巧上的重要性，人们发现唯象理论比微观理论更有用处”。金兹堡也因此获得了 2003 年的诺贝尔物理学奖。

就在刚刚过去的 2021 年，诺贝尔物理学奖被授予“对我们理解复杂系统做出了开创性贡献”的三位科学家<sup>[55]</sup>。其中，意大利物理学家乔治·帕利西（Giorgio Parisi）因为“发现从原子到行星尺度的物理系统中的无序和涨落的相互作用”而分享了一半的奖金。同样，Parisi 所做的工作，就是针对自旋玻璃中复本对称性破缺开展了序参量描述，刻画了多个自旋态的自组织关系。

由此可以看出，朗道的序参量思想在方法论方面有着重要的价值，并一直被后人所延续和发展。如前文所述的“协同学”，就是将朗道的序参量思想引入复杂系统的研究，实现了对自组织的定量刻画。那么，对于更为复杂的系统，面对写不出演化方程或者无法求解方程的情况，该如何开展复杂系统的研究呢？让我们把目光投向近期湍研究所取得的突破。

## 2.2 结构系综理论的突破

湍流是自然界中一个典型的复杂系统，具有开放性、大自由度、多尺度、非线性、高度自组织和整体有序化等特征<sup>[56]</sup>。湍流问题，就是要从看似随机、混乱无序的流体运动中，揭示出整体有序性中蕴含的规律。正因为其复杂性，一百多年来没有得到解决，号称“经典统计力学中最后一个尚未解决的难题”。

余振苏教授在多年的湍流等复杂系统的研究中，拓展了朗道的“序参量”的思想，建立了“结构系综理论（Structural Ensemble Dynamics, SED）”<sup>[57,58]</sup>。

结构系综理论（SED）认为，系统有序性的来源，来自于宏观的对称破缺。因为发生对称破缺以后，在宏观对称性的约束下，系统会涌现出相似结构。而相似结构，一定满足某种变换不变性。序参量，就是此时系统的统计结构在宏观所展现出来的相似变量。所以，即使不知道方程或者无法求解方程，但是通过对称性分析，也可以从实测数据中去把握相似变量，也就是序参量。根据序变量满足的变换不变性，就可以写出该问题物理解的解析表达。

具体来说，对于壁湍流系统，由于强剪切导致了速度脉动，从而导致出现随机拉伸对称性的破缺，产生若干个沿壁面法向的涡旋结构系综，造成雷诺应力、湍动能的反常标度行为。由于脉动速度的出现，湍流平均运动方程不封闭，从而导致严格解析无法获

得，微扰方法也失效。

结构系综理论（SED）从对称性分析出发，认为真正受到拉伸对称性约束的是湍流结构系综，而结构系综的特征尺度才是正确的相似变量，而不是传统的湍流相似性研究所关注的平均速度或者雷诺应力张量本身。因此，结构系综理论（SED）将序参量的概念拓展为“序函数”，定义了“长度序函数”来描述流动涡结构系综的特征尺度：

$$\ell_{12} = \frac{\sqrt{\langle -u'v' \rangle}}{\partial U / \partial y} \quad (2.1)$$

固壁对结构系综提供了宏观的约束，该约束的对称变化群是拉伸群，所以壁湍流中的约束被称为拉伸对称性约束，因此长度序函数随空间坐标满足幂次律（拉伸不变性的解）。进一步，结构系综理论（SED）的一个重要发现是，相邻结构系综之间应力长函数满足普适的幂次律跃迁形式，称为广义拉伸对称性原理。这一原理使得理论上能解析地将空间不同的结构系综连接起来，实现能够覆盖湍流边界层垂向全域的多层结构的解析表述，从而形成描述湍流边界层平均速度和湍动能全剖面的 SED 解析理论。如图 2.1 所示，从壁面向外依次形成的粘性底层、缓冲层、对数层、外区主流层便是逐次对称破缺的结果，物理上形成一个四层结构。这样一个四层结构覆盖了垂直壁面的全流域。四个应力长的四层结构函数表达为：

$$\ell_{12}^{+SED} = \ell_0 \left( \frac{y^+}{9.7} \right)^{3/2} \left( 1 + \left( \frac{y^+}{9.7} \right)^4 \right)^{0.5/4} \left( 1 + \left( \frac{y^+}{y_{buf}} \right)^4 \right)^{r_2/4} \frac{(1-r^4)}{4(1-r)} \quad (2.2)$$

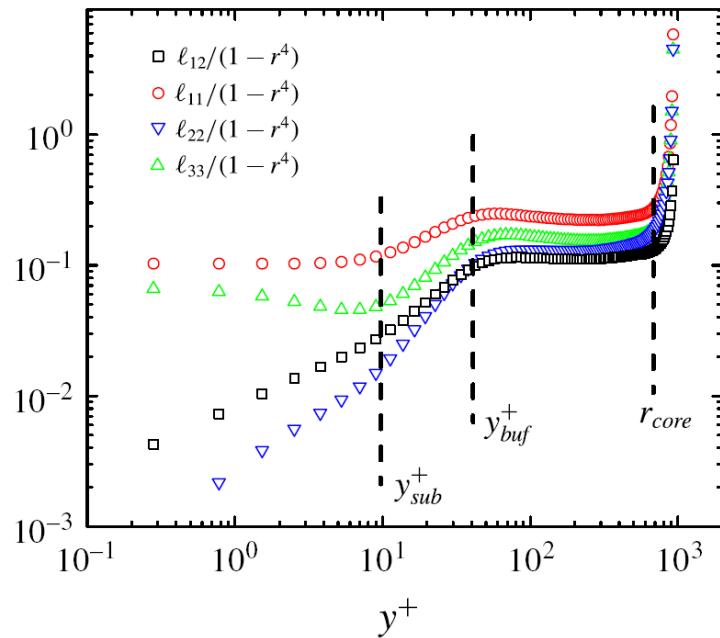


图 2.1 壁湍流的长度序函数的四层结构<sup>[58]</sup>

近年来,结构系综理论(SED)已经成功地运用到刻画规范边界层、可压缩边界层、RB热对流等多种典型流动系统,以及大气表面层等真实流动系统,并且成功开发了新型湍流模型,应用于工程实践<sup>[57,58]</sup>。

李荣博士<sup>[59]</sup>将结构系综理论(SED)应用于高温超导,发现与湍流中的涡旋相似,赝能隙态中涨落的密度波序(相干长度为1-2个周期)的关键物理参数为其周期(约1-3纳米)。这些短程序产生的有效势不断发生涨落,导致对载流子的倒逆散射,进而决定了赝能隙、其起始温度以及散射速率与密度波序周期成平方反比关系。这个能量法则得到了100多个样品的验证,解释了赝能隙以及它的起始温度随掺杂增加单调减小的原因,揭示了赝能隙来源会从轻度掺杂区的自旋密度波涨落转变为中间掺杂区的电荷密度波的涨落,回答了Goldman问题,即不同临界电阻是由于临界载流子浓度的差异。该研究第一次揭示了铜氧化物中的超导、赝能隙以及奇异金属相都具有的普适性规律,表明介观涨落序(无论是超导、密度波还是涡旋)是刻画强关联电子系统的关键,值得进一步深入研究。

综上所述,结构系综理论(SED)发展了“序参量”的思想,将单一的相空间的对称破缺,拓展到实空间的多对称破缺,因而序参量的数学描述中出现了多维多层次的自相似律和自组织原理。这对于刻画具有多维多层次的对称性破缺的相似律和自组织原理的复杂系统,例如人体复杂系统,具有重要意义。

## 2.3 人体序参量理论探索

根据第一章介绍的人体复杂系统科学理论,人体是一个一元二面多维多层次的复杂系统,存在着多维多层次的对称破缺的相似律和自组织原理,所以可以通过多维多层次的序参量来进行刻画。通过对于二类具体的人体系统的研究,本文尝试发展一套研究人体复杂系统的人体序参量理论。

### 2.3.1 研究思路

在“一元二面多维多层次”的人体观下,人体复杂系统研究的重点在于通过把握多个“二面”的互动规律,从而实现对于人体自组织中心的多维多层次的刻画。那么,如何把握多个“二面”的互动规律呢?

余振苏教授提出,将朗道的“序参量”思想引入到人体复杂系统的研究中来。每一对提炼出的人体系统的“二面”,实际上是对人体系统的某种对称破缺的刻画。因此,针对具体系统,开展对称性分析,就可以凝练出背后隐藏的关键序参量。进一步,根据多种约束条件,可以直接写出序参量满足的解析表达式,构建出序参量模型,由此实现对人体复杂系统不同的有序态的定量刻画。据此,本文提出了基于“序参量”刻画人体

自组织中心，探索人体复杂系统研究的新思路：

第一，综合前人经验和研究成果，开展对称性分析，提炼主导序参量，针对具体的生命现象，建立描述系统宏观有序态的唯象模型。

第二，将唯象的系统论模型与定量的微观动力系统方法联合应用，以实现对人体生命现象进行从宏观到微观的刻画。

第三，在全面刻画人体生命现象数据的基础上，不断提炼其中蕴含普遍的科学原理，发现普适常数。而这些原理的提炼，又促成了新的科学问题的提出。由此，就形成了一个“多层表述、逐级定量、多次迭代，逐步近似”的研究框架。

### 2.3.2 研究方法

根据上述人体序参量理论，开展具体的研究方法，主要由以下三步组成：

第一，针对具体的研究系统，开展“一元二面”的对称性分析，明确“二面”背后的对称破缺机制，从而确定出主导这一过程的序参量。对于人体复杂系统而言，存在着多维多层次的对称破缺，这使得无法一下子就明确背后究竟发生了什么对称破缺。但是，“一元二面”的方法论对于我们从以往的丰富的研究成果中直接凝练出重要的序参量，提供了重要的支撑。例如，本文在分析人体健康系统时发现，以往的研究注重于刻画人体的当前状态，而其当前状态距离“常态”的“偏离度”才是主导其健康水平的序参量。由此，实现了基于脉搏波对人体健康水平的定量刻画。

第二，根据系统满足的约束条件——比如守恒条件、边界条件等，可以直接写出序参量所满足的解析表达式。这是整个研究中最为精彩的一步，同样也是最有挑战性的一步。这需要对所研究系统的演化图像，有更为清楚的认识。而“一元二面多层次”的认识论，为此提供了具体的方法途径。例如，本文在刻画新冠疫情演化时发现，根据抗疫图像的系统论分析，病毒增长、社会抗疫力量都符合早期由零开始指数增长、晚期饱和为一常数的演化特征，由此直接确定出了疫情演化和死亡率演化背后所满足的统一的 Logistic 的演化行为。

第三，根据序参量的解析表达式，就可以进一步刻画实测数据。在这个过程中，探索参数的演化规律，发现背后的普适规律。复杂系统的研究中总是存在所谓“特例”的情况，这是因为存在着新的对称破缺机制。因此，在研究中需要拥抱这些特例，发现背后隐藏的规律。例如，本文在刻画各地区死亡率与床位占有率之间的幂律关系时，发现武汉市距离幂律曲线较远。进一步分析发现，如果考虑了武汉市修建方舱医院增加了床位数后，就满足模型的刻画结果了。

上述三步，在实际工作中，也不是一蹴而就的，需要来回不断验证和迭代。这样，就构成了一个“多次迭代、逐步近似”的人体复杂系统研究的方法路线。

### 2.3.3 研究进展与特点

人体序参量理论的构建，属于一项原始创新性的研究。前人也曾提出过要将“序参量”的思想引入到人体复杂系统的研究中来。例如，郭蕾等人提出“阴阳”是人体系统的序参量<sup>[60,61]</sup>；吴邦惠等提出“气”是人体系统的序参量<sup>[62]</sup>；陶必修提出“信息”和“气血”是人体系统的序参量<sup>[63]</sup>。但是，这些讨论仍停留在理论假设层面，并没有提出具体的数理模型以及开展实验验证。

而吕嘉戈和马蔼乃<sup>[64]</sup>基于相似理论，通过量纲分析，构建了人体气血系统、筋骨系统、肌肉神经系统、皮毛结构、脏腑系统、“奇恒”之腑的数学模型，显示了通过寻找序参量开展人体复杂系统建模的有效性。

张陆昊等人<sup>[65]</sup>将人体序参量理论应用于人体系统内部新冠病毒模拟的工作中，基于六个关键的序参量，构建了一个简洁的三层次的新冠肺炎动力学模型，从而突破了传统由于囊括了大量的免疫组分导致的“维数灾难”的挑战。通过将模拟结果与数百名患者的实时临床数据进行对比，该研究实现了对死亡和存活患者的免疫应答差异以及与T细胞的关系的量化。该研究发现，T细胞免疫在所有患者的病毒清除功能中都占据主导地位(88.8%)，但是这种贡献从轻症(96.7%)到重症患者(80.4%)会有明显的下降，并伴随着T细胞数目的减少。在抑制炎症方面，细胞因子的抑制速率和组织修复速率会也随着T细胞数量的减小而减小，并且在死亡者中几乎为零，由此揭示了T细胞免疫功能在减少新冠肺炎患者死亡中的关键作用。这一工作受到了多方面的关注，显示了上述方法论在处理人体复杂系统问题的有效性。

综上所述，上述人体序参量理论的研究，是一个唯象的定量研究。这一方法的重要特点是：

第一，在尚不清楚复杂的微观作用机制的情况下，可以对系统的宏观演化，开展从定性到定量的刻画，从而发现宏观层次的规律——而这又成为更深入的微观机制研究的基础。这对于研究人体这样具有多个维度、多个层次的复杂系统，是十分必要的。

第二，在尚不清楚发生的对称破缺机制的情况下，可以直接写出序参量所满足的解析表达式。由此就可以直接开展对不同演化情况的定量刻画，发现背后存在的普适规律。众所周知，对于真实的复杂系统，是很难写出符合真实演化情况的方程的——即使可以，也很难求解。例如，上述介绍的湍流难题。而相比湍流，人体系统更加复杂。如果能够直接写出典型生理过程的典型结构——例如本文提出的高斯波包解，那么对于进一步研究背后所反映的人体规律，是十分有帮助的。

第三，通过刻画大量的多种演化场景下的数据，来提炼出系统演化的自组织机制。针对不同的研究对象，提炼不同的序参量，因此这一理论具有较大的包容性。而更重要的是，对于同一类信号，其构建序参量模型的方法又是普适的。例如，本文构建的刻画

人体健康偏离度的模型方法，既适用于刻画脉搏波信号，同时也适用于刻画其他人体波动信号（如 GDV 信号）。

## 2.4 应用于人体复杂系统

本文从群体的疫情演化到个体的健康分析，开展了人体序参量理论的研究。由此构成了一幅从群体数量变化到人体健康变化再到人体内部脉搏波的脉动量涨落的研究图景。

### 2.4.1 群体：新冠疫情演化系统

已经持续了两年多的新型冠状病毒肺炎（以下简称“新冠”）大流行（COVID-19），对于人体复杂系统的深入研究，提供了丰富的研究数据。

面对新型冠状病毒（SARS-CoV-2），有着不同文化背景和不同干预和医疗能力的国家和地区，采取了不同的防疫政策，这些多方面的环境因素作用在具有不同的人口学特征的个体上，最终在疫情演化行为上呈现出了巨大的差异。

更为重要的是，在这个过程中，来自数学、物理、化学、生物学、医学、计算机、公共政策和社会学等多方面的科学家开展了深入的交叉合作，形成了对新冠病毒和疫情演化等多方面的丰富认识。正如《自然》期刊所说，“在科学史上，人们第一次在全世界范围内收集和分享了如此多的信息”<sup>[66]</sup>。针对这些来自于不同方面、不同层次的认识，进行有效的从定性到定量的综合集成，凝练出疫情演化的一般规律，正是人体复杂系统研究所要面对的重要课题。

具体来说，对于新冠传播的复杂系统，我们看到的是各类人群数字的变化，但实际背后是病毒增长与环境约束之间的对称性变化。刚开始时，病毒增长是以指数的速率来增长，但是由于环境的约束，病毒的增长速率受到限制，最终会趋于一个饱和值。同样，社会的干预力量也符合这样一个过程，最开始社会的干预力量较弱，但是随着疫情的爆发，干预力量的增长速率呈指数增长，最终因为疫情得到控制，也会趋于饱和。所以，这个过程中，存在着一个基本的统计结构。假设病毒的增长速率  $x'/x$ ，由于受到环境约束，基于自身规模而线性衰减  $(\lambda - ax)$ ，那么就有  $x' = x(\lambda - ax)$ 。这就是著名的种群增长模型，即 Logistic 模型<sup>[67]</sup>。这就是新冠演化系统中的基本统计结构。

本文将这一结构首先应用于刻画疫情的传播，发现疫情传播中的三大过程，分别由三个序参量刻画——即针对易感人群是否被传染的“隔离度”、针对轻症人群是否被治愈的“治愈度”和针对重症人群是否被抢救的“抢救度”，都符合 Logistic 的形式。根据人数满足的对称性约束，成功构建了新的传染病动力学模型。并且在刻画国内外多种真实疫情演化场景中，进行了模型的验证。这部分研究的具体内容，详见第三章。

进一步，我们将这一结构应用于刻画新冠死亡率的演化，发现全球 52 个国家在第一波疫情中的全周期死亡率都符合这一统计结构的规律。并且通过构建包含了疫情、人口、社会的序参量的相空间，发现了疫情在早期、晚期阶段分别满足的普适规律。这部分研究的具体内容，详见第四章。

#### 2.4.2 个体：人体脉搏波系统

在我国抗击新冠肺炎疫情的阻击战中，中医药抗疫成了中国方案的“亮点”，彰显了中国智慧。根据《抗击新冠肺炎疫情的中国行动》白皮书显示，“中医医疗队整建制接管定点医院若干重症病区和方舱医院。……中医药参与救治确诊病例的占比达到 92%。湖北省确诊病例中医药使用率和总有效率超过 90%。”<sup>[68]</sup>国家中医药管理局设立应急专项，对 10 个省市 1261 名服用“清肺排毒汤”的患者临床观察显示，没有 1 例转为重型或危重型<sup>[69]</sup>。这些成绩足以说明，中医药把握住了人体复杂系统的重要规律。所以，即使面对新型的病毒和复杂的疫情，中医药也能够发挥可靠的作用。

实际上，在绵延几千年的中华文明历史中，中华儿女一直都是依靠中医药来保障健康的。在这个过程中，中医积累了丰富的临床诊疗经验，形成了对人体复杂性的较为系统的认识。在这些宝贵的知识和经验的基础上，针对客观数据，开展从定性到定量的综合集成，正是人体复杂系统研究的重要内容。

具体来说，脉诊在中医诊疗中具有辨别病情、阐述病机、指导治疗和推断预后等重要作用，说明脉搏波信号中蕴含着丰富的人体系统的信息，所以我们选择从脉搏波信号切入研究人体复杂系统。对于人体脉搏波系统，我们看到的形态各异的脉搏波压力波形。但实际上它背后反映了传输过程中发生的多次对称性变化。根据前人研究，神经信号、心电、脉搏波等多种人体波动信号都可以由高斯波包的形式来很好地进行刻画。这显示了人体生理的波动信号背后遵循着统一的对称性。虽然目前我们还不能很清楚地确定这个对称性是什么，但我们猜想它所产生的统计结构，就是高斯波包形式。通过研究发现，脉搏波满足的运动方程的解可由四个高斯波包组成，并且成功地拟合了人体多种脉象的脉搏波。四个高斯波包对应了人体脉搏波在传播过程中发生多次对称破缺的过程，这对于理解人体生理和病理过程，有重要的意义。这部分具体内容，详见第五章。

进一步，对于人体健康系统，我们仍然从对称性出发，认为人体在长期演化作用下发生对称破缺形成了个性化的“常态”，同时又不断发生新的对称破缺，造成与“常态”的偏离。因此，与“常态”之间的“偏离度”，才是刻画人体宏观有序度的关键序参量。偏离越远，有序度越低，则健康水平越差。偏离越近，有序度越高，则健康水平越高。在此基础上，我们提出“偏离度”与“健康指数”之间满足的基本统计结构为最简单的负线性关系，结果显示无论是个体常态的健康指数还是个体日变的健康指数，都符合

这一线性关系。这一模型将中医脉诊理论的“知常达变”思想进行了数理化的定量表达，突破了传统基于大数平均统计研究无法进行个性化精确刻画的局限。这部分的具体内容，详见第六章。

综上所述，人体序参量理论是一项正在研究和发展中的理论。本文通过将其基本思想和方法应用于群体的“新冠疫情演化系统”和个体的“脉搏波系统”中，发现了其中都存在着多次对称破缺的过程，而已有的研究往往只关注于某一个对称破缺过程。因此，通过凝练多次对称破缺过程的序参量和确定约束条件，本文分别构建了序参量模型，精确地刻画了相关数据，并从中总结出普适规律和关键常数，展示了人体序参量理论的有效性和可靠性。同时，这一研究也丰富了人体序参量理论的内涵，发展了人体复杂系统研究的方法。



## 第三章 应用于新冠疫情传播系统

2019 年末，湖北省武汉市爆发了新型冠状病毒肺炎疫情（以下简称新冠疫情，COVID-19）。2020 年 3 月 11 日，世界卫生组织（WHO）宣布新冠疫情进入全球大流行状态。在大流行持续两年后，截止 2022 年 5 月 1 日，全球累计感染人数已经超过了 5.1 亿人，死亡 620 万人以上。

这次罕见的疫情引起全世界人民的关注，国际卫生组织定期公布全球疫情变化的详细信息，为开展细致的传染病动力学建模，揭示疫情演化的科学规律提供了条件。正确的数学模型应该能够预测未来疫情演化的趋势，对于政策制定者实施科学的抗疫措施，起到重要的作用<sup>[70]</sup>。本章从人体序参量理论出发，建立了一个新的基于对称性的传染病传播动力学模型，将其应用于定量刻画国内外不同特征地区、不同流行波次、不同干预政策的疫情演化行为，揭示了当代大国抗击疫情中，三类国家干预措施在其中起到的主导力量。

### 3.1 新冠疫情传播的复杂性

根据《自然》期刊的报道，保守估计，仅 2020 年发表的新冠研究论文就超过了 10 万篇。而在这些论文的主题中，关于模拟疫情传播的论文在同年的前七个月中一直是数量最多的<sup>[71]</sup>。这一方面反映了人们对于新冠疫情建模研究揭示疫情演化规律的巨大需求，同时也反映出新冠疫情建模的复杂性。

本节首先简要介绍了新冠疫情传播建模的研究进展，进一步分析了新冠疫情演化的复杂性，由此提出了当前研究的突破点。

#### 3.1.1 疫情传播的建模研究

人们针对特定疾病已经开发了许多模拟疾病传播的模型，其背后的数学思想基本上是一致的，最早可以追溯到著名的法国数学家丹尼尔·伯努利 (Daniel Bernoulli, 1700-1782)。在他生活的时代，正是“天花”这一流行病的高发时期，来自于东方的人痘接种术可以治愈一部分天花患者，但仍然有较高的死亡率。伯努利通过将人群分为感染者和未感染者，且一部分感染者会通过接种而治愈，由此建立了数理方程。并且通过计算，他得出结论：接种人痘虽然有危险，但仍能延长三年存活期<sup>[72]</sup>。今天来看，这一模型虽然十分简单，但却包含了流行病传播建模的最核心思想。

在此基础上，苏格兰科学家 Kermack 和 McKendrick 于 1927 年提出了著名的仓室模型 (compartment model)<sup>[73]</sup>。这一模型将人群分为：易感者 S (Susceptible，即易被

感染的健康人)、感染者 I (Infectious, 即已经被感染的人) 和康复者 (Recovered, 即感染后已经康复的人)。这就是著名的 SIR 模型。

对于像新冠肺炎这样的传染病，易感者在感染病毒后存在一定的潜伏期，因此特别划分出新的一类人群——暴露者 E (Exposed)，即接触过感染者但还没有传染性的群体。假设人群是均匀混合的，则同一仓室中的个体将以相同的转变率转移到下一个仓室。这就是目前应用最广泛的 SEIR 模型。

在疫情早期，有一些研究团队选择直接运用 SEIR 模型，讨论了新冠疫情的演化情况<sup>[74,75]</sup>。而更多的研究结合新冠疫情的特点，对 SEIR 模型主要进行了以下三个方面的改进：

第一类，通过增加新的仓室，来模拟更真实的人群分布<sup>[76-82]</sup>。例如，文献<sup>[77]</sup>增加了被限制的人群 (confined)，由此可以通过讨论不同的解除“被限制的人群”的方式，模拟了不同的解封政策对于疫情演化的影响。

第二类，通过将原来仓室之间不变的转换率调整为时间依赖的函数，来模拟更为真实的疫情演化<sup>[83-90]</sup>。例如，文献<sup>[84]</sup>将治愈度和死亡率修改为时间依赖的函数，来更好地刻画治愈度的提高和死亡率的降低，有其他文献<sup>[85-89]</sup>直接引用该模型或在此模型的基础上进行了改进。

第三类，通过将各群体再进行细分，以及考虑气象变化<sup>[91]</sup>、交通网络<sup>[92]</sup>、人口流动数据<sup>[93]</sup>等，来模拟更为真实的人群混合情况。以至于细分到个体，开展基于个体 (agent-based) 的建模<sup>[94]</sup>。但是这类模型需要大量的社会接触数据，来确定相关的模型参数，而这却又往往难以直接获得<sup>[70]</sup>。

### 3.1.2 已有研究的局限

实际上，通过将以上部分 SEIR 模型的预测<sup>[80,81]</sup>与我国新冠疫情数据进行对比，可以显著地看到，模型预测的早期增长和峰值是较为合理的，但是预测的疫情衰退显著地慢于实际的疫情衰退。其根本原因在于——SEIR 模型假定核心参数（如传染速率、治愈速率、死亡速率）为常数。然而，实际疫情数据显示，这三大转换速率从疫情的早期到后期通常具有 1—3 个数量级的变化，反映了当代大国战疫中国家干预能力的大幅度提升及其对疫情演化的重要影响<sup>[84]</sup>。并且各国的国情不同，采取了不同的防疫政策，以及病毒的不断变异<sup>[95]</sup>等等因素交织在一起，使得全球疫情演化呈现出十分复杂的演化模式。因此，基于简单假设的 SEIR 模型，无法刻画当代各国抗击新冠疫情的真实系统<sup>[91]</sup>。

针对这一局限性，如上所述，目前的研究主要是通过增加仓室和调整转换速率的方法，来对模型进行改进。这些改进虽然使得模型能够刻画好一部分实际数据，但同时又带了更大的不确定性。

首先，对于增加仓室的改进模型，由于模型中只有确诊人群、治愈人群和死亡人群的数据是被公开统计报道的，可以约束相应的模型参数，而更多的参数只能通过经验或假设来确定。当引进越多的仓室时，这意味着将会有更多的自由参数，这给模型的稳定性和可靠性带来了挑战<sup>[91,96]</sup>。

其次，对于转换率的改进，往往具有较强的随意性，并且未给出充分的物理依据。例如，Peng 等人<sup>[84]</sup>提出的改进的 SEIR 模型将治愈速率和死亡速率修改为时间依赖的函数，并被很多文献所引用<sup>[85-88]</sup>。又如 CHRynet<sup>[89]</sup>就采用了多种函数形式来拟合治愈速率和死亡速率，并从中选取拟合效果最好的函数形式，来刻画疫情演化模式。这些改进具有较强的随意性，虽然能够刻画好一部分疫情数据，但是由于不同地区采用了不同的函数形式，对于应用模型来揭示疫情演化中的普适规律提出了挑战。

综上，人们已经意识到转换速率的变化对于此次疫情演化至关重要，它反映了国家干预措施的影响<sup>[84]</sup>，但仅仅是作为一种“经验参数”而一笔带过，没有引起学者们足够的关注和讨论。因此，随着大流行的持续，亟需建立一个新的符合真实疫情演化的动力学系统，来揭示国家干预措施对疫情演化的作用规律，并对国家防疫政策制定提供更深入的科学依据。

本章从上文介绍的人体序参量理论出发，对新冠疫情演化这个群体动力学系统开展了对称性分析，发现除了 SEIR 模型所刻画的病毒主导的对称破缺过程外，还存在由国家干预措施主导的对称破缺过程。进而，通过确定这一过程中的主导疫情演化的三大序参量及其 Logistic 演化行为，即针对易感人群的隔离度 ( $G$ )、针对轻症的治愈度 ( $H$ ) 和针对重症的抢救度 ( $R$ )，构建了一套新的基于动力学模型系统，揭示了疫情演化中存在的“序参量-转换速率-群体数量”这三个层次相互耦合的复杂系统自组织动力学机制。进一步，本文将这一模型应用于定量刻画不同特征地区、不同流行波次和不同干预政策等真实的疫情演化场景，展示了模型在定量刻画各级政府干预措施方面的有效性。这对于政策制定者开展科可持续的抗疫措施，尽量减少疫情对人们和社会的影响，具有广阔的实践意义。

## 3.2 新冠疫情传播的多层次对称性刻画

我们将疫情爆发的过程看成是一个对称破缺过程，于是可以采用上文介绍的人体序参量理论，对疫情传播这个复杂系统开展对称性分析。即要回答，在新冠疫情传播的过程中，人群系统究竟发生了什么“对称破缺”的过程呢？

### 3.2.1 病毒出现后的对称性刻画

首先，是由新病毒在人群中传播所造成的对称破缺。在没有新病毒时，人群系统中

的不同个体之间的接触没有什么差别，即具有一种对称性。当新病毒（如 SARS-CoV-2）出现后，人与人之间的接触具有了传染和被传染的方向性，即原来人群系统的对称性发生了破缺。进一步，人群系统还可以细分为易感者、潜伏者、康复者等，这些人群之间的接触也具有了新的有序性。而不同人群之间的转换率，可以刻画这种有序性。如图 3.1 所示，易感者与感染者接触后，会以一个转换率  $a$  转换为暴露者。当  $a = 0$  时，说明群体之间不会发生转换，疫情显然就不会传播，人群系统具有对称性。而当  $a = 1$  时，说明只要接触感染者，所有的易感者就会 100% 被感染。同理，暴露者会以转换率  $b$  成为感染者，进而会以转换率  $c$  成为恢复者。进一步写出人数满足的守恒方程，就是著名的 SEIR 模型。因此，我们从对称性的角度解释了 SEIR 模型之所以在刻画疫情传播有效的原因：即它抓住了新病毒出现所导致的人群对称破缺的过程，而人群之间的转换率就是刻画有序性的序参量。

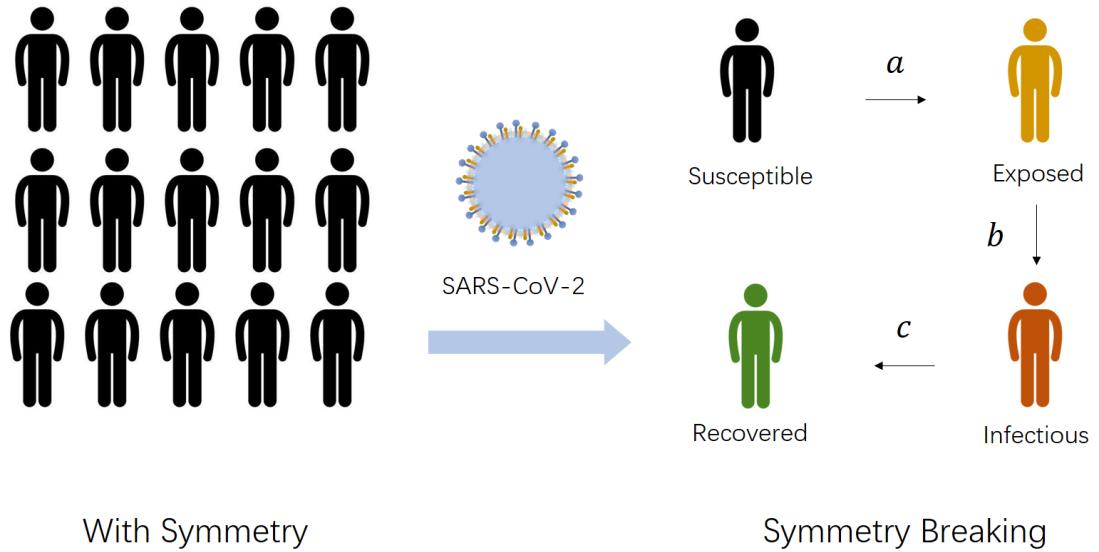


图 3.1 新病毒的出现导致人群发生对称破缺。黑色为易感者，浅黄色为暴露者，橘红色为感染者，绿色为康复者。

由于满足人数守恒的约束，我们就可以写出人群之间的转换。例如，易感者 ( $S_v$ ) 与感染者 ( $I_v$ ) 接触后，会以一个转换率  $a$  转换为暴露者 ( $E_v$ )，即：

$$\frac{dS_u}{dt} = -a * S_v * I_v \quad (3.1)$$

同时，暴露者 ( $E_v$ ) 会以转换率  $b$  成为感染者 ( $I_v$ )，则有：

$$\frac{dE_v}{dt} = a * S_v * I_v - b * E_v \quad (3.2)$$

同时，感染者 ( $I_v$ ) 会以转换率  $\gamma$  成为恢复者 ( $R_v$ )，则有

$$\frac{dI_v}{dt} = b * E_v - \gamma * I_v \quad (3.3)$$

$$\frac{dR_v}{dt} = \gamma * I_v \quad (3.4)$$

上式 3.1-3.4 就是著名的 SEIR 模型。由此，我们从人体序参量理论出发，通过对称性分析，提炼出人群之间的“转换率”这个序参量，并进一步根据序参量满足的数学结构和人数守恒的约束，写出了 SEIR 模型。这就从理论上解释了 SEIR 模型之所以一直以来在刻画疫情有效的原因：它抓住了新病毒出现后人群对此破缺的过程。

### 3.2.2 国家干预后的对称性刻画

当国家实施大规模干预措施后，上述各类人群之间的转换过程会受到相应措施的影响作用，而发生从“病毒自由传播”到“被限制的传播”的对称破缺，如图 3.2 所示。我们可以将国家干预措施主要分为三大类，即针对易感人群的隔离措施、针对轻症的治疗措施和针对重症的抢救措施。

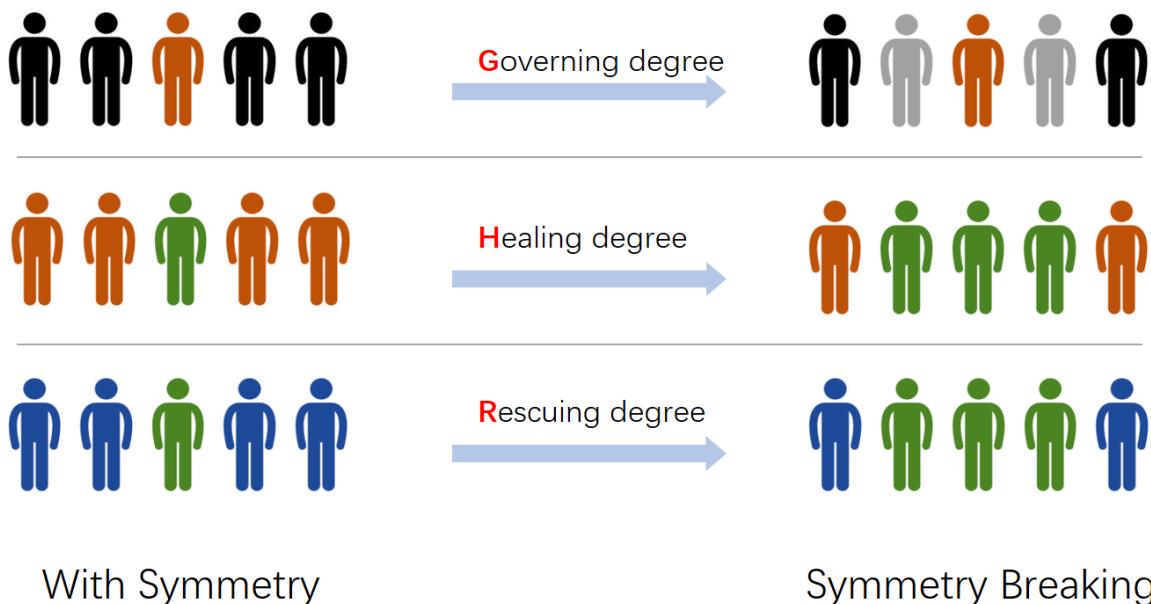


图 3.2 国家三类干预措施导致相应人群进一步发生对称破缺。黑色为易感者，灰色为受保护的易感者，橘红色为感染者，绿色为康复者，蓝色为重症患者。

### 隔离度 G

第一类是对易感人群开展的隔离措施。隔离措施包括了封城、限制社交距离社会隔离措施，也包括个体戴口罩、勤洗手等自我隔离措施。这些措施使得病毒不能在人群

中自由传播，而是只能向某一部分人传播，例如密切接触者，即产生了一种新的“序”。

因此，我们提炼出了“隔离度”( $G$ )来刻画隔离措施的力度。当 $G = 0$ 时，社会或个人没有开展隔离措施，病毒在人群中可以自由传播。当 $G = 1$ 时，说明社会或个人开展了严格的隔离措施，则病毒无法在人群中传播。

## 治愈度 $H$

第二类是针对轻症人群开展的治疗措施，包括增加医护人员和轻症床位（如修建方舱医院）、补充医疗资源、提高医护能力、推广标准治疗方案等等，这些措施增加了医疗力量，使得不同程度的患者的治疗和转归有了不同的方向，即产生了一种新的“序”。

因此，我们提炼出了“治愈度”( $H$ )来刻画治疗措施的力度。当 $H=0$ 时，国家没有开展治疗措施，轻症人群得不到救治。当 $H=1$ 时，说明国家开展了有力的治愈措施，所有被感染的人群都能够被救治。

## 抢救度 $R$

第三类是针对重症患者的抢救措施，包括增加 ICU 床位、增加生命支持仪器等，以此来使得一部分重症患者能够被有效抢救而转归为轻症病例，即产生了一种新的“序”。

因此，我们提炼出了“抢救度”( $R$ )来刻画抢救措施的力度。当 $R = 0$ 时，国家没有开展抢救措施，重症人群被救治的几率较低。当 $R = 1$ 时，说明国家开展了有力的抢救措施，所有重症都能够被救治，则不会再出现死亡病例。

综上所述，我们通过运用人体序参量理论的方法，对疫情传播系统开展了对称性分析，发现存在着由新病毒和国家干预措施所导致群体系统发生了多层次的对称破缺过程。对于新病毒出现导致的对称破缺过程，我们从序参量的角度，对传统的 SEIR 传染病动力学模型进行了理解。而对于 SEIR 模型所没有考虑的国家干预导致的对称破缺过程，我们提炼出了隔离度( $G$ )、治愈度( $H$ )和抢救度( $R$ )三大序参量。下面，我们将在此基础上行，构建一个定量刻画国家干预的传染病动力学模型——GHR 模型。

### 3.3 新冠疫情传播的 GHR 模型

#### 3.3.1 序参量的解析表达

根据人体序参量理论，在提炼出表征系统有序性的序参量后，进一步通过确定出序参量背后满足的数学结构，结合约束条件或边界条件，就可以直接写出序参量的解析表达。那么，隔离度( $G$ )、治愈度( $H$ )和抢救度( $R$ )这三个关键序参量满足什么数学结构呢？

在疫情早期，新病毒处于自由传播的状态。这时，国家还没有开展干预措施，所以隔离度 ( $G$ )、治愈度 ( $H$ ) 和抢救度 ( $R$ ) 等于 0。随着疫情的爆发，这三类措施开始呈指数增长。但是，由于社会资源的限制，国家干预措施不会一直增长下去，所以这三类干预措施的增长速率会放缓，最终达到一个饱和值。

因此，我们假设这三类干预措施的相对增长速度，是随着干预力度的增加而线性衰减，则有：

$$\frac{dG}{dt} = \gamma G \left(1 - \frac{G}{g_0}\right) \quad (3.5)$$

$$\frac{dH}{dt} = \alpha H \left(1 - \frac{H}{h_0}\right) \quad (3.6)$$

$$\frac{dR}{dt} = \mu R \left(1 - \frac{R}{r_0}\right) \quad (3.7)$$

其中， $g_0$  为饱和的隔离度 ( $0 \leq G \leq g_0$ )， $h_0$  为饱和的治愈速率 ( $0 \leq H \leq h_0$ )， $r_0$  为饱和的抢救度 ( $0 \leq R \leq r_0$ )。 $\gamma$  是隔离度  $G$  的增长速率， $\alpha$  是治愈度  $H$  的增长速率， $\mu$  是抢救度  $R$  的增长速率。

这就是著名的 Logistic 函数，也被称为“生长曲线”或“S 曲线”。这一函数经常被用于描述生物系统的演化，例如微生物的增长，生物种群的增长，人体重量、身高随年龄的增长，人体光子辐射信号随年龄的变化等<sup>[97]</sup>。值得注意的是，这一基本的函数形式能够被广泛地应用于光子、微生物、人体、群体等多个尺度的生物系统的演化。这本身就不平凡，很可反映了 Logistic 函数抓住了生物系统演化中的普适的统计结构。在后面章节，本文运用实际报告的疫情数据，证明了上述三类国家干预力量的增长，也符合这一基本的统计结构。

### 3.3.2 GHR 模型

进一步，根据上文分析，在疫情早期，国家干预措施的力度为 0 时，此时的传染速率 ( $V_I$ ) 和死亡速率 ( $V_D$ ) 分别为病毒自由传播的恒定值  $c_0$  和  $s_0$ 。而由于病人治疗康复需要一定的时间，此时的治愈速率 ( $V_C$ ) 几乎为 0。随着国家干预措施力度增加，传染速率 ( $V_I$ ) 和死亡速率 ( $V_D$ ) 最终衰减到 0，而治愈速率则增长到一个饱和值。所以，假设序参量与转换速率之间为最简单的线性关系，则有：

$$V_I(t) = c_0(1 - G(t)) \quad (3.8)$$

$$V_C(t) = H(t) \quad (3.9)$$

$$V_D(t) = s_0(1 - R(t)) \quad (3.10)$$

另外，定义传染速率 ( $V_I$ )、治愈速率 ( $V_C$ ) 和死亡速率 ( $V_D$ ) 为：

$$V_I(t) = \frac{\Delta I(t)}{Q(t)} \quad (3.11)$$

$$V_C(t) = \frac{\Delta C(t)}{Q(t)} \quad (3.12)$$

$$V_D(t) = \frac{\Delta D(t)}{Q(t)} \quad (3.13)$$

上式中的现有确诊病例数  $Q(t)$ 、新增感染病例  $\Delta I$ 、新增治愈病例  $\Delta C$  和新增死亡病例  $\Delta D$  之间满足人数守恒律，则有：

$$\frac{dQ}{dt} = \Delta I - \Delta C - \Delta D \quad (3.14)$$

由此，3.5 式—3.14 式，就构成了一个新的我们称之为 GHR 的传染病动力学模型。相比传统的 SEIR 模型，GHR 模型的核心不仅仅关注各类人群的变化，而更着重于刻画国家干预措施对于疫情演化的影响。如图 3.3 所示，国家干预措施的变化，将导致隔离度 ( $G$ )、治愈度 ( $H$ ) 和抢救度 ( $R$ ) 这三个关键序参量的变化，进一步体现在相应的传染速率 ( $V_I$ 、治愈速率 ( $V_C$ ) 和死亡速率 ( $V_D$ ) 的变化上。而最终，呈现在人们眼前的，则是每天报告的感染病例  $I(t)$ 、治愈病例  $C(t)$  和死亡病例  $D(t)$  等病例数据的变化。由此，就构成了一幅“序参量-转换速率-人群演化”这三个层次之间相互耦合的疫情演化图像，揭示了疫情复杂系统自组织的动力学机制。

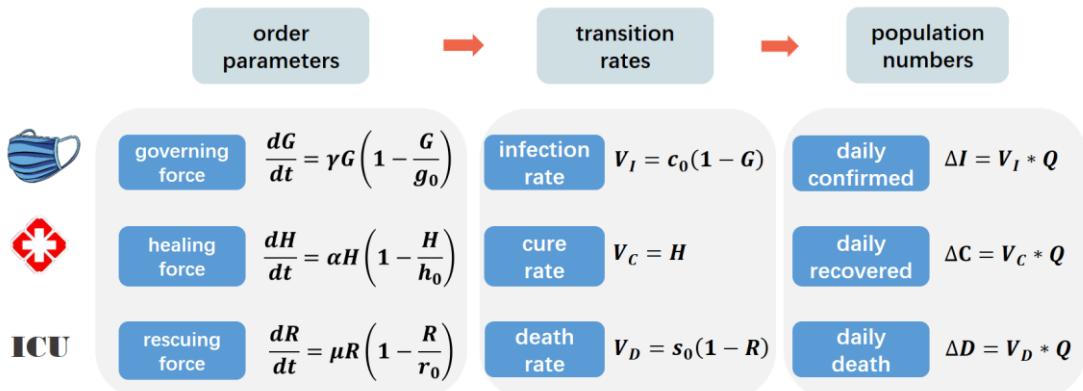


图 3.3 GHR 模型示意图

### 3.3.3 模型求解及参数意义

三个序参量为著名的 Logistic 函数（见 (3.5) — (3.7) 式），存在解析解：

$$G(t) = \frac{g_0}{1 + e^{-\gamma(t-t_g)}} \quad (3.15)$$

$$H(t) = \frac{h_0}{1 + e^{-\alpha(t-t_h)}} \quad (3.16)$$

$$R(t) = \frac{r_0}{1 + e^{-\mu(t-t_r)}} \quad (3.17)$$

其中， $t_g$ 、 $t_h$ 、 $t_r$ 分别为 $G$ 、 $H$ 、 $R$ 三个 Logistic 函数的中点所对应的特征时间。

下面，我们来对模型参数进行梳理和总结，以隔离度 $G$ 为例：

当 $t \gg t_g > 0$ 时，3.15 式可以写成：

$$G(t) \approx g_0 \quad (3.18)$$

因此， $g_0$ 表示在疫情晚期的饱和隔离度。

当 $t = t_g$ 时，3.15 式可以写成：

$$G(t_g) \approx \frac{g_0}{2} \quad (3.19)$$

因此， $t_g$ 表示隔离度达到饱和值的一半时的特征时间。

当 $t \approx 0$ 时， $G(t) \ll g_0$ ，3.5 式可写成：

$$\gamma \approx \frac{G(t)'}{G(t)} \quad (3.20)$$

这是一个指数增长函数，所以 $\gamma$ 表示疫情早期时的隔离度的指数增长速率。

当 $t \approx 0$ 时， $G(t) \approx 0$ ，3.8 式可以写成：

$$V_I(t) \approx c_0 \quad (3.21)$$

因此， $c_0$ 表示疫情早期病毒自由传播时的传染速率。

由此，我们就解释了对隔离度所涉及的模型参数的物理含义。同理，我们可以得到治愈度和抢救度的模型参数的物理意义。总结起来，GHR 模型包含了 8 个模型参数，根据三类序参量，可以分为三组，它们的含义如下：

(1)  $c_0$ 代表疫情早期时，病毒在无隔离措施时的传染速率； $\gamma$ 是隔离度( $G$ )的增长速率，决定了早期的隔离度( $G$ )的指数增长的快慢； $g_0$ 是最终的饱和隔离度。

(2)  $\alpha$ 是治愈度( $H$ )的增长速率，决定了早期的治愈度( $H$ )的指数增长的快慢； $h_0$ 是最终的饱和治愈速率。

(3)  $s_0$ 代表疫情早期时，没有国家干预时的死亡速率； $\mu$ 是早期的抢救度( $R$ )的增长速率，决定了早期的抢救度( $R$ )的指数增长的快慢； $r_0$ 是最终的饱和抢救度。

特别地，当  $t = 1$  时，治愈度（11 式）可以写为：

$$e^{-\alpha(1-t_h)} = \frac{h_0}{H(1)} - 1 \quad (3.22)$$

进一步可以写为：

$$\alpha = \frac{1}{t_h - 1} \ln \left( \frac{h_0}{H(1)} - 1 \right) \quad (3.23)$$

由于  $h_0 \gg H(1)$ ，上式可以简化为：

$$\alpha \approx \frac{-1}{t_h - 1} \ln(H(1)) + \frac{1}{t_h - 1} \ln(h_0) \quad (3.24)$$

因此，当  $1/(t_h - 1)$  和  $\ln(h_0)$  近似为常数时，治愈度的初值  $H(1)$  与其增长速度  $\alpha$  之间为对数律关系。本文在后面章节基于中国大陆 23 个省份的疫情数据，对这一些关系进行了验证。

### 3.3.4 差分方程系统

前文给出了疫情发展的连续微分方程系统。但报告数据实际上是以“天”为单位而呈离散性的。因此，刻画疫情的传播应该采用以“天”为步长的差分方程。

当天新增的确诊（治愈/死亡）病例数取决于此前已经存在的确诊病例，及其在当日以一定的传染（治愈/死亡）速率。因此，我们提出对前述的连续微分方程转变为差分方程表达：

$$Q(i) - Q(i - 1) = \Delta I(i) - \Delta C(i) - \Delta D(i) \quad (3.25)$$

$$\Delta I(i) = c_0(1 - G(i)) * Q(i - 1) \quad (3.26)$$

$$\Delta C(i) = H(i) * Q(i - 1) \quad (3.27)$$

$$\Delta D(i) = s_0(1 - R(i)) * Q(i - 1) \quad (3.28)$$

其中， $i$  代表疫情中的某一天， $\Delta I$ 、 $\Delta C$  和  $\Delta D$  分别代表当日新增的确诊、治愈和死亡病例数。三类序参量隔离度 ( $G$ )、治愈度 ( $H$ ) 和抢救度 ( $R$ ) 则分别取相应的 Logistic 函数解析解，即 (3.15) – (3.17) 式。

在 GHR 模型中， $Q$  代表“现有确诊”的人数；同时， $\Delta I$ 、 $\Delta R$ 、 $\Delta D$  对应于新增的确诊人数、治愈人数和死亡人数；而它们对时间的积分则对应于累计的确诊人数、治愈人数和死亡人数。因此，GHR 模型通过 11 个参数（包括 8 个模型参数和 3 个特征时间参数）刻画了七类病例数的演化行为。这种刻画在数据完整的情况下，参数的冗余度和不确定性通常比较小。因此，在确定合理的参数确定方法后，可获得可靠的模型参数，

并给出对疫情演化动力学的定量描述。

### 3.3.5 参数反演算法

针对一个地区，我们采用官方公布的累计确诊、累计治愈和累计死亡三类病例数据，来反演 8 个方程参数 $\{c_0, \gamma, g_0, \alpha, h_0, s_0, \mu, r_0\}$ 和 3 个特征时间参数 $\{t_g, t_h, t_e\}$ 。

多参数的准确确定，也是本研究的一大特色。参数可以分三组，顺次独立进行参数反演。具体来说，先反演隔离度的控制参数 $\{c_0, \gamma, g_0, t_g\}$ ，再反演治愈度的控制参数 $\{\alpha, h_0, t_h\}$ ，最后是死亡率的控制参数 $\{s_0, \mu, r_0, t_r\}$ 。具体流程见图 3.4。

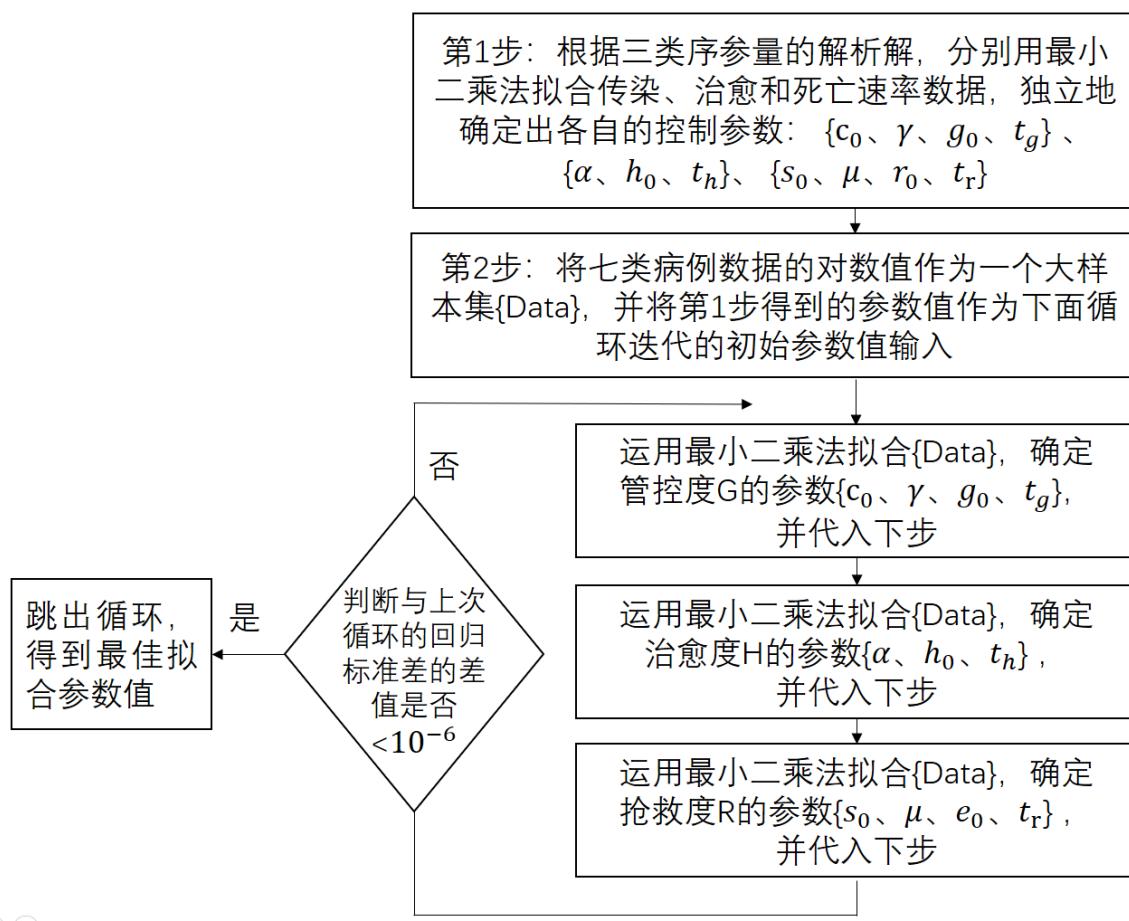


图 3.4 参数反演的流程图。

### 独立刻画速率获得参数近似值

第一步是分别对三类序参量的解析解（3.15—3.17 式）开展最小二乘拟合，独立获得相应的控制参数。这里我们采用了 Matlab® 软件内置的非线性拟合函数——lsqcurvefit 进行最小二乘拟合。以湖北省的情况为例，我们通过最小二乘拟合三类序参量的对数

值，分别得到各自的参数近似值。

但是，由于实际报导的疫情数据通常会在个别时间点上出现较奇异的行为，例如，由于诊断标准的改变，某天忽然新增大量确诊病例，等等。这些奇异点导致实际报导的累计值总是大于由（光滑变化的）速率所积分得到的值。这些奇异点的影响可以通过模型参数调整来反映。调整参数后所预测的（每天变化的）速率比报导的速率略大，原因就是吸收了少数日期的奇异点的贡献。这就是由下述第二步所实现的。

### 联合刻画病例数获得最佳参数值

第二步，通过刻画七类病例数据（现有/新增/累计确诊、治愈和死亡病例）来获得模型的最佳参数值。

首先，将七类病例数据的对数值作为一个大样本集，计算第一步所有样本点的误差的平方和对应的回归标准差<sup>[98]</sup>：

$$\hat{\sigma} = \sqrt{\frac{1}{n-p-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3.29)$$

其中， $y_i$ 是第*i*日的观测数据， $\hat{y}_i$ 对应的预言值， $p=1$ 是自变量的数目， $n$ 是总的数据量。

其次，通过依次对三类序参量的三组参数进行最小二乘拟合，每次都将更新后的相应参数代入下一轮。通过比较三步结束时的 $\hat{\sigma}$ 与三步拟合之前的初始 $\hat{\sigma}$ 的差异，如果小于 $10^{-6}$ ，则得到参数结果，否则继续如上的循环，最终得到最小回归标准差 $\hat{\sigma}_{min}$ 、最优模型参数值以及病例数的最佳拟合。

### 参数搜索获得不确定度

考虑到数据涨落的客观存在性，需要对每个参数的不确定度进行估计：首先，固定其中10个参数为最佳参数值，让另外1个参数在最佳值的附近进行扫描，计算参数变化带来的回归标准差 $\hat{\sigma}$ 的变化曲线，如图3.5。为了确保较好的实际拟合效果，我们以回归标准差达到最小回归标准差 $\hat{\sigma}_{min}$ 的1.1倍时的参数范围，作为不确定度的上界，即：

$$\hat{\sigma}_{ub} = 1.1 * \hat{\sigma}_{min} \quad (3.30)$$

图3.5中的黑色虚线代表了 $\hat{\sigma}_{ub}$ ，它与回归标准差曲线的最外侧的两个交点即刻画了参数的上下界，从而给出参数的不确定度，见表3.1。

在参数反演中，我们充分利用了已知的解析近似解来获得参数的近似值，围绕着近似值搜寻最佳参数值和开展不确定性分析。在表3.1中，我们展示了由此得到的浙江

省、湖北省和意大利的模型参数值和初值。

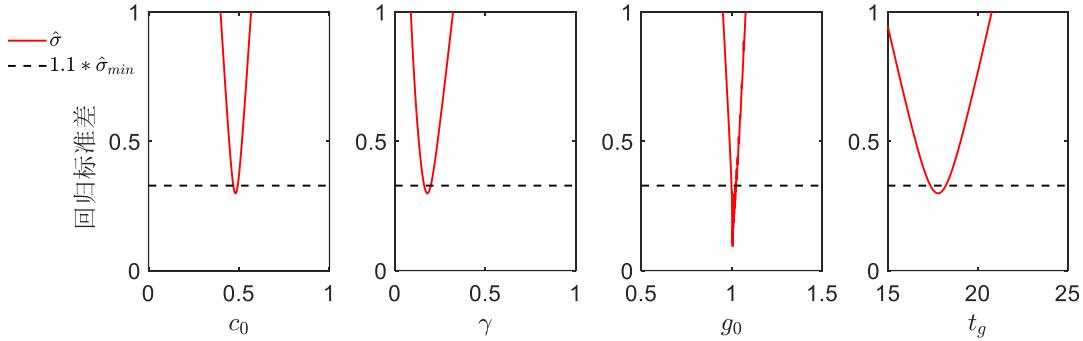


图 3.5 湖北省隔离度  $G$  的参数不确定度。其中红色实线是当该参数值变化时的回归标准差，黑色虚线是  $\hat{\sigma}_{ub}$  的值，两条线交点的横坐标即为参数取值范围，见表 3.1。

表 3.1 参数的定义和各地区的模型参数值和初值

序参量	含义	湖北	意大利
$c_0$	初始传染速率	$0.48 \pm 0.01$	$0.48 \pm 0.02$
隔离度 $G$	$\gamma$ 隔离度的增长速率	$0.18 \pm 0.02$	$0.10 \pm 0.02$
控制参数	$g_0$ 饱和的隔离度	$1.00 \pm 0.00$	$0.98 \pm 0.01$
	$t_g$ 隔离度减半时间	$17.8 \pm 0.4$	$17 \pm 1.0$
治愈度 $H$	$\alpha$ 治愈度的增长速率	$0.08 \pm 0.01$	$0.03 \pm 0.01$
	$h_0$ 饱和治愈度	$0.15 \pm 0.01$	$0.09 \pm 0.04$
	$t_h$ 治愈度达半时间	$50 \pm 2$	$96 \pm 20$
抢救度 $R$	$s_0$ 初始死亡速率	$0.03 \pm 0.01$	$0.03 \pm 0.01$
	$\mu$ 抢救度的增长速率	$0.17 \pm 0.06$	$0.50 \pm 0.50$
	$r_0$ 饱和的抢救度	$0.95 \pm 0.02$	$0.80 \pm 0.10$
	$t_r$ 抢救度达半时间	$9 \pm 5$	$32 \pm 18$

## 3.4 定量刻画国家干预力度

### 3.4.1 疫情爆发中心地区的比较

湖北省和意大利分别属于新冠第一波疫情在亚洲和欧洲的爆发地区，具有相似的动力学演化特征，并且二者在人口、面积、纬度等方面也具有较好的可比性。因此，我们运用 GHR 模型来定量地比较湖北省和意大利的疫情演化特征，以揭示背后国家干预力量的主导作用。

需要说明的是，意大利所报告的数据不具有疫情早期演化的特征，即不拥有早期传染速率 $c_0$ 的平台，导致 $c_0$ 无法被准确确定。于是，我们假设意大利（在未知的早期）与湖北省的初始传染速率一致。确实，通过数据分析发现两者在确诊人数达到 100 人以后，其传染速率的演化非常相似（从几乎相同的初值 0.43 和 0.42，到大概 18 天后速率减半），如图 3.6d 所示。这也对我们的假设提供了自洽的支撑。另外，意大利每日新增治愈数在达到 1000 的量级后仍会出现新增为 0 的情况，说明其早期的治愈病例的报导不规范，因此导致其早期的治愈速率数据的涨落太大（见图 3.6e）。所以，我们在确定模型参数时，选择暂时忽略这些早期治愈数据，而选取 36 天后的数据来拟合确定出治愈速率参数。

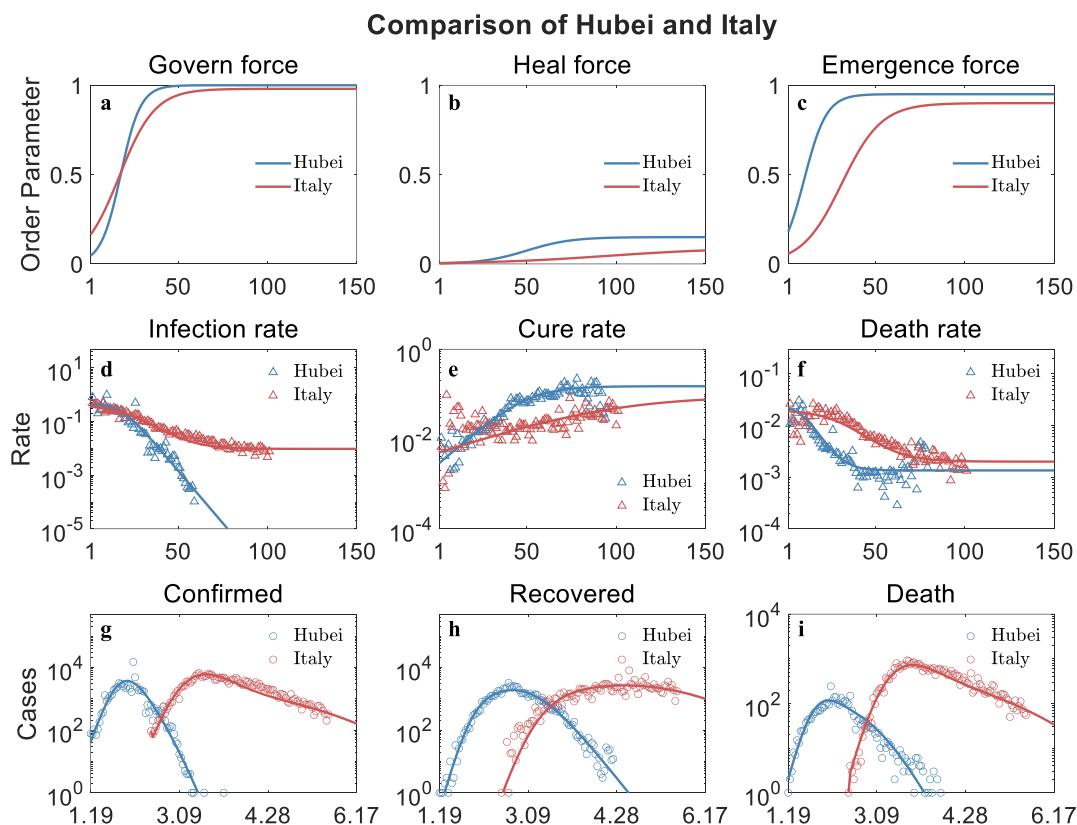


图 3.6 湖北省和意大利的疫情演化对比。红色和蓝色圆圈分别是湖北省卫健委和约翰·霍普金斯大学疫情实时监测系统收集的数据（截止于 2020 年 6 月 2 日），三角形是由此计算出的速率数据，实线则是 GHR 模型模拟结果。模拟参数见表 3.1。以现有确诊达到 100 例时为第 1 天来进行比较。a-b 为序参量的演化结果，其中 a 为隔离度、b 为治愈度、c 为抢救度；d-f 为三类人群转换速率的演化结果，其中 d 为传染速率、e 为治愈速率、f 为死亡速率；g-i 为群体数量的演化结果，其中 g 为新增确诊病例、h 为新增治愈病例、i 为新增死亡病例。

首先，3.6 图展示了一个完整的“序参量”（图 3.6a—3.6c）“转换速率”（图 3.6d—3.6f）和“病例数量”（图 3.6g—3.6i）三个层次相互耦合的演化图像。以传染过程为例，

隔离度 $G$ 在 0-1 之间呈现出 Logistic 形式的增长行为（图 3.6a），由此导致了相应的传染速率发生 2-3 个数量级的降低（图 3.6d），最后体现在病例数量上，为新增确诊病例的“增长-衰减”的演化行为（图 3.6g）。同样的，治愈过程和死亡过程也都很好地符合这一规律，由此验证了模型的有效性。

其次，我们发现在疫情早期，三类转换速率的初值都十分接近（见图 3.6d—3.6f），初始的传染速率约为 0.5，治愈速率约为 0，而死亡速率约为 0.03。由于在第一波疫情时两地的毒株是一样的，因此接近的转换速率验证了 GHR 模型所提出的疫情早期的演化行为是由病毒自由传播所主导的假设。更重要的是，这些关键常数反映了毒株本身在传染力、致死力等方面的能力。未来值得通过与其他变异株或病毒进行比较，来确定它们在这些方面的特点。

最后，GHR 模型证明了国家干预力量对疫情走向的主导作用。以传染过程为例，如图 3.6a 显示，意大利隔离度的初始值 $G(1)$ 实际上比湖北省还要高，这与意大利得到湖北爆发疫情的警示有关，但湖北省的隔离度的增长速率 $\gamma$ 是意大利的 1.8 倍（见表 3.1），这取决于文化的差异和政府干预力度的不同，例如，根据新闻报道，意大利的隔离封锁政策启动较晚，在现有确诊病例达到 6387 人（3 月 8 日）时才开始封锁伦巴第大区，而湖北省在现有确诊 494 人时即对武汉封城。因此，这就导致在大约 20 天后，湖北省的隔离度就超过了意大利，并最终导致了意大利的累计确诊病例（23.3 万）是湖北省（6.8 万）的 3.4 倍。由此可见，GHR 模型在定量刻画不同地区社会抗疫力度上取得了重要进展。

更重要的是，GHR 模型对于科学判断解封时间和预测疫情二次爆发的可能性，提供了定量的依据。如图 3.6a 所示，湖北省的隔离度 $G$ 大约在 65 天后就接近于 100%，传染速率降至 0，由此实现了新增确诊病例的“清零”。而事实上，武汉就是在此继续封城了 14 天后成功地实现了解封。与之相比，意大利的饱和隔离度 $g_0$ 最终只增长至 98%，导致传染速率下降到 1% 左右时就进入到了“平台期”，这意味着意大利的传染过程可能会较长期的存在。根据媒体报道，意大利于 2020 年 5 月 4 日开始逐步解封，而此时的隔离度 $G$ 仅为 97.4%，最终造成了疫情的二次爆发。因此，饱和隔离度 $g_0$ 是判断疫情走向的重要参数。

同样，疫情传播中的治愈过程和抢救过程，也都体现了相应的国家干预力量的关键性作用。例如，如果在其他参数不变的情况下，模型模拟的结果显示，意大利治愈度的增长速率 $\alpha$ （0.03）达到湖北省的 $\alpha$ （0.08）时，将有大约 7.5 万人免于感染，约 1.5 万人免于死亡。同样，如果意大利对危重症的抢救度的增长速率 $\mu$ （0.09）达到湖北省的 $\mu$ （0.16）时，最终将会避免约 1.4 万人死亡。

综上所述，GHR 模型定量地揭示了疫情爆发中心在疫情早期的由病毒主导的过程，以及之后的由国家干预主导的过程，为定量评估抗疫成效提供了支撑。

### 3.4.2 23个省市的两类治愈度演化模式

相比于湖北和意大利等疫情爆发中心地区，对在良好隔离下的广大疫情外溢地区的动力学特征进行建模分析，可以揭示国家干预措施中的普适规律。

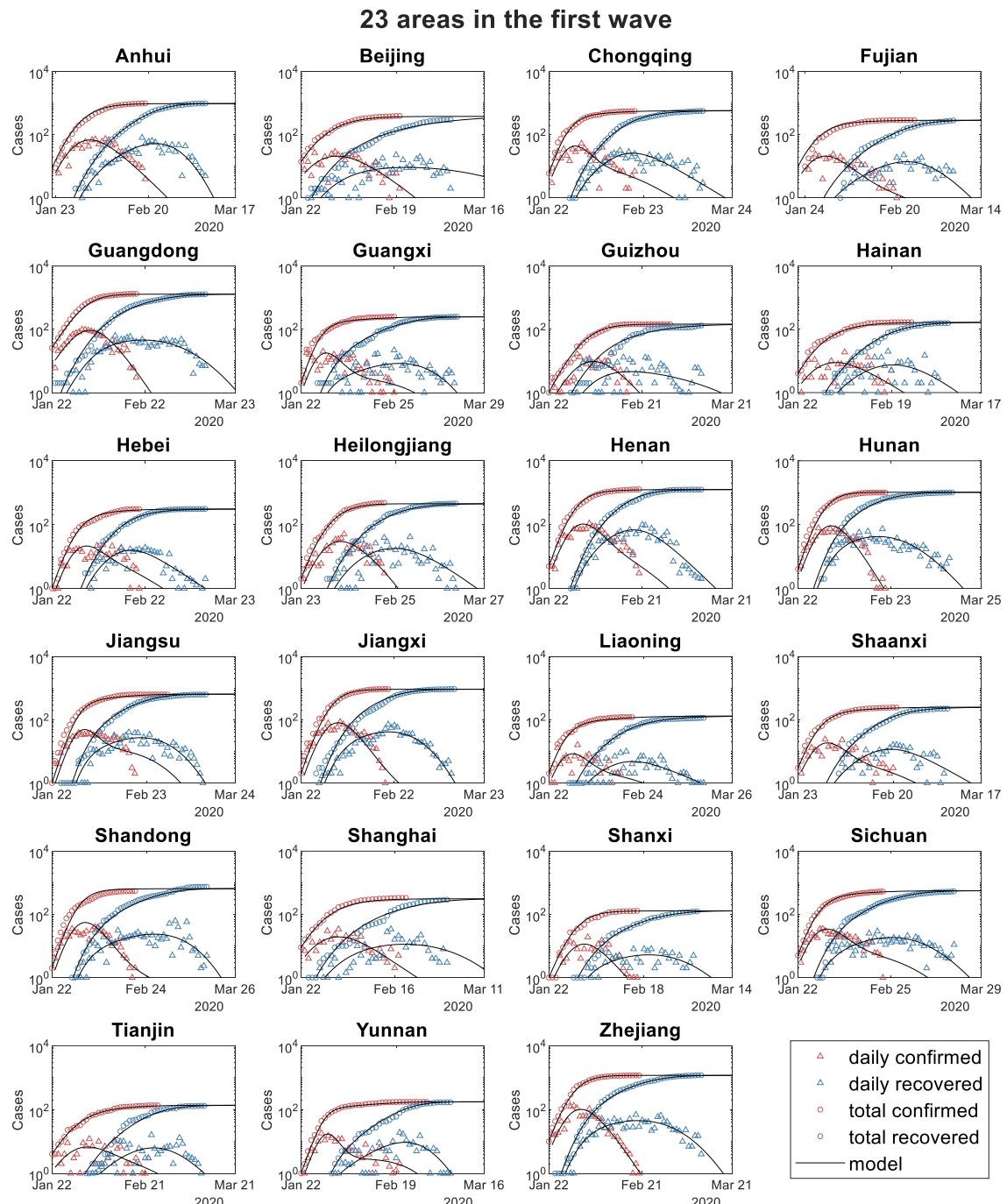


图 3.7 全国第一波疫情中的 23 个省份的模型刻画结果。圆圈为来自于约翰·霍普金斯大学疫情实时监测系统收集的累计数据（截止于 2020 年 4 月 1 日），三角形是由此计算出的每日新增数据，实线则是 GHR 模型的模拟结果。

这些地区由于疫情被迅速地进行控制，医疗系统能够较为良好地运转，从而死亡率普遍较低。例如，在第一波疫情中，除湖北以外的 30 个疫情外溢省份的平均死亡人数仅为 4 人。针对这一特点，GHR 模型可以通过将初始死亡速率  $s_0$  设置为 0 来对模型进行约化，从而实现对这些地区的疫情传播开展解析刻画。考虑到统计稳定性的要求，本文排除了在第一波疫情中累计感染病例小于 100 例的 7 个省份，对剩下的 23 个省份的疫情数据成功地进行了模拟分析，如图 3.7 所示。

首先，与疫情爆发地区形成鲜明对比的是，疫情外溢地区的政府响应速度很快，隔离度  $G$  和治愈度  $H$  的平均增速都是湖北省的 1.5 倍，这使得隔离度和治愈度的达半时间比湖北省分别提前了 11.3 天和 6.6 天（见表 3.2）。从而，这些地区取得了平均  $27 \pm 5$  天清零，感染  $528 \pm 395$  人的令人印象深刻的抗疫成绩。

表 3.2 中国大陆 23 个疫情外溢省份的模型参数的平均值与方差

Govern force ( $G$ )			Heal force ( $H$ )		
$\gamma$	$g_0$	$t_g$	$\alpha$	$h_0$	$t_h$
$0.27 \pm 0.09$	$99.0\% \pm 1.0\%$	$6.5 \pm 2.7$	$0.12 \pm 0.05$	$0.32 \pm 0.14$	$43.4 \pm 18.9$

进一步，我们发现在由治愈度的初始值  $H(1)$  及其增速  $\alpha$  所组成的相空间中，23 个省份满足一个良好的对数关系（见图 3.8a）。根据 3.24 式，这一关系的普适性反映了各省份治愈度的达半时间十分接近，同时这也有力地证明了各省份的治愈度  $H$  具有 Logistic 的演化模式。

有趣的是，这些省份同时还自动聚集为两个明显的团簇（见图 3.8a）。第一类是以河北等地区为代表的治愈度演化的模式（N=11）。它们在疫情初期的治愈度  $H(1)$  比较低，其平均值仅为第二类的三分之一，但治愈度的平均增速  $\alpha$  却是第二类的 2 倍（见表 3.2）。这说明，这些地区的治愈度  $H$  在疫情持续期间得到了政府干预力量的积极响应。而以北京为代表的第二类演化模式（N=12）的地区在地理上展示出明显的分布特征，它们基本上都是我国东部、南部或经济相对较为发达的省份（直辖市）（见图 3.8b）。这些地区的医疗资源相对充足，因此在疫情初期就有相对较高的治愈度  $H(1)$ ，并且以一个相对较低的增长速度，也能够完成对疫情的良好控制。进一步，我们还发现这两类演化模式还体现在关键群体演化的“持续时间”上。首先，考虑到疫情数据在早期和晚期存在“长尾效应”，我们根据峰值前后发生一个数量级的上升和下降的时间范围，来确定出相应人群演化的持续时间（见图 3.8c）。我们也曾用这一方法发现了全球 52 个国家的死亡人群演化的持续时间与 ICU 非存活者的持续时间之间具有普适的线性律（见第四章）<sup>[99]</sup>。在这里，我们发现 23 个省份的康复病例和在院病例的演化持续时间之间具有

普适的线性律（见图 3.8d）。根据 3.14 式，在院病例的演化主要是由确诊病例和康复病例两部分所决定的。而这一普适的线性律则进一步揭示了，对于疫情外溢地区，康复病例是在院病例演化的主导因素。这一结果说明，在良好的社会隔离度和治愈度的支持下，疫情外溢地区的医疗系统能够良好运转，感染病例主要是以轻症病例为主，从而能够得到及时的治疗而很快被治愈出院（见图 3.7）。因此，提高对轻症的治愈度能够有效地缩短疫情的持续时间，尽量地减轻医疗系统的压力。

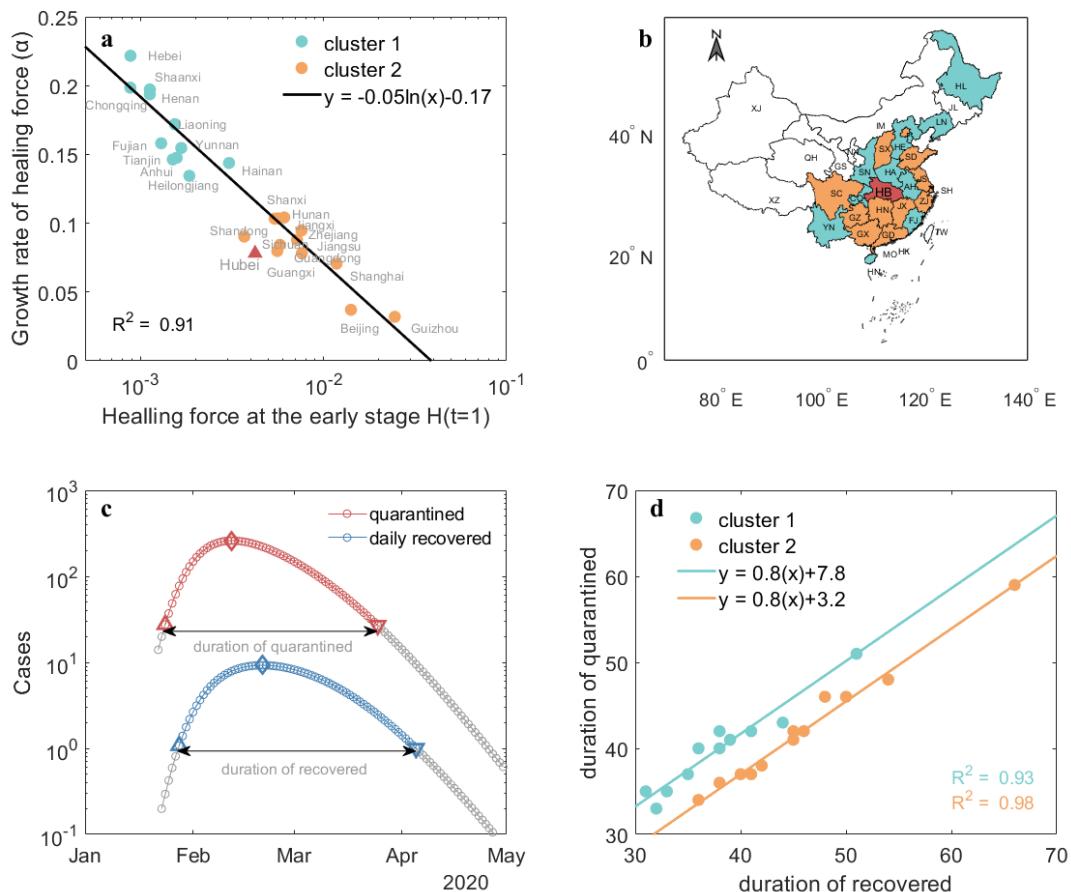


图 3.8 两类治愈度的演化模式。a, b, d 中的浅蓝色代表第一类模式，橘黄色代表第 2 类模式。  
 a 在治愈度相空间中的两类演化模式。b 两类演化模式在地理上的分布情况。c 对于康复病例（蓝色点）和在院病例（红色点）的持续时间的定义（以北京市为例）。d 两类演化模式的康复病例和在院病例的线性关系。

值得注意的是，上述两类治愈度演化的模式具有各自的线性标度（见图 3.8d），这同时也再次支持了上述两类治愈度演化模式的划分结果。二者的斜率十分接近（0.8），而区别仅在于截距不同，即第一类地区的康复病例演化的持续时间都小于在院病例演化的持续时间，而第二类则恰好相反。进一步，我们从表 3.3 中可以看出，两类模式在确诊病例和在院病例的演化时间上都十分接近，但第二类的康复病例的演化时间比第

一类多了 6 天，确诊人数多了 31%（189 人）。这很可能是因为第二类地区的人口聚集度高，所以确诊人数稍多，从而康复病例演化的持续时间更长一些。

表 3.3 两类演化模式的关键统计量的对比

Cluster	Number	$H(1)$	$\alpha$	Confirmed	Duration of confirmed	Duration of recovered	Duration of quarantined
1	11	0.003 ±0.004	0.16±0.05	428±358	28±4	40±10	42±7
2	12	0.009 ±0.006	0.08±0.02	617±406	26±5	46±8	42±7

### 3.4.3 北京市在两波疫情中的成功经验

我们还将 GHR 模型应用于定量刻画不同流行波次的演化特征，以揭示政府的干预措施在其中所起作用的微妙变化。这里，我们选取了北京市在 2020 年 1 月爆发的第一波疫情，以及在同年 6 月爆发的第二波疫情。这两波疫情相隔的时间较短，并且都属于同一种早期的新冠病毒毒株，因此具有一定的可比性。所不同的是，北京市的第一波疫情是由于湖北省的病例迁移而导致的局部外溢式的模式。而第二波疫情则是由于新发地农产品市场聚集扩散导致的中心爆发式的模式。并且由于第二波疫情的爆发地环境与当初武汉华南海鲜市场十分相似，引起了人们的关切。我们运用 GHR 模型对这两波疫情都进行了很好的模拟，结果如图 3.9 所示。

首先，第一波疫情的隔离度和治愈度的初始值，都大于第二波疫情（见图 3.9a, b）。这与湖北爆发疫情后，北京迅速启动了“重大突发公共卫生事件一级响应”的事实一致。而第二波疫情是当地爆发的疫情，在政府干预前，病毒已经自由传播了一段时间，所以初期的干预力度较低。但是，第二波疫情的隔离度的增长速率  $\gamma$  是第一波疫情的 2.3 倍，所以在第 9 天时隔离度就超过了第一波疫情，并迅速接近于饱和值 100%（见表 3.4）。同样，第二波的治愈度的增长速度  $\alpha$  是第一波的 8.5 倍，从而在 20 天左右时就接近于饱和值（见图 3.9b）。这说明北京市在第二波疫情时，在短时间内就调动起了远超过在第一波疫情时的强度的干预力量。

其次，我们注意到两波疫情的初始传染速率和治愈速率都十分接近（见图 3.9c, d），这与两波疫情都是同一种毒株的事实一致。随着疫情发展，我们看到在国家干预力量的作用下，第二波疫情的传染速率更为迅速地降低了近 3 个数量级，从而比第一波疫情提前 5 天就成功地实现了新增病例的清零（见表 3.4）。

最终，第二波疫情的累计感染人数为 335 人，甚至比第一波疫情的 395 人还要少，并且死亡病例为 0。这已经成为了国家干预措施下的典型成功案例。然而，如果第二波

疫情仍然采取第一波疫情的隔离度度，模型预测最终感染人数将超过 2.2 万人，这将对医疗系统带来超额的负荷。因此，上述结果充分地展示了，即使对于疫情爆发地区，通过迅速地提高国家干预措施的力度，也能够将疫情控制在较小的规模内，尽量地减少死亡病例。

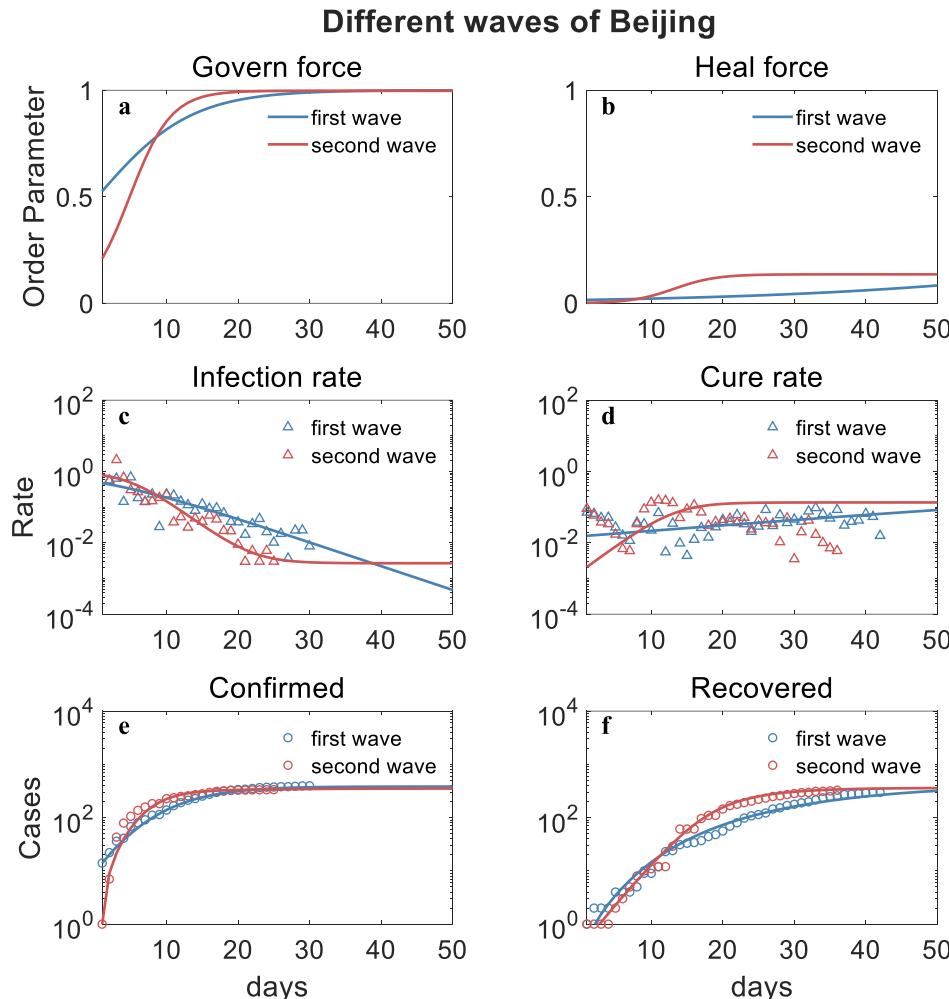


图 3.9 北京市不同流行波次的疫情演化特征比较。红色和蓝色圆圈分别是湖北省卫健委和约翰·霍普金斯大学疫情实时监测系统收集的数据（截止于 2020 年 7 月 2 日），三角形是由此计算出的速率数据，实线则是 GHR 模型模拟结果。模拟参数见表 3.4。a-b 为序参量的演化结果，其中 a 为隔离度、b 为治愈度；c-d 为三类人群转换速率的演化结果，其中 c 为传染速率、d 为治愈速率；e-f 为群体数量的演化结果，其中 e 为新增确诊病例、f 为新增治愈病例。

表 3.4 两类演化模式的关键参数和统计量的对比

Wave	$\gamma$	$G(1)$	$\alpha$	Confirmed	Zero time
1	0.15	0.53	0.04	395	25
2	0.35	0.21	0.34	335	30

### 3.4.4 定量模拟动态清零政策

2021年底，面对传染力倍增的奥密克戎（Omicron）变异株，中国政府提出了“动态清零”的疫情隔离策略，即通过采取综合防控措施来快速扑灭疫情。这一政策在2022年的吉林、深圳等地的疫情隔离中，显示了令人印象深刻的结果。但同时，对于上海、北京这样千万级人口的一线城市，如何更好地平衡疫情防控和经济活动，是这一政策深入有效实施所面临的挑战。对此，我们提出，需要定量地监测和评估“动态清零”政策的隔离效果，从而保证有足够的调整空间来灵活地采取综合的干预措施。因此，我们运用GHR模型对北京市在2022年4月22日爆发的第三波疫情数据，讨论了在不同隔离度下的疫情演化，结果如图3.10所示。

首先，从实际报道的病例数据来看，北京市第三波疫情的演化呈现出与前两波疫情截然不同的特点。例如，本轮新增确诊病例（不包含“无症状感染者”）在不到一周就超过了50例，突破了此前历史最高值（44例）。并且这一高峰期不像前两波疫情被迅速遏制而走向衰退，而是进入到一段持续20余天的平台波动期，呈现出“双峰”等更为复杂的演化特征。传染速率则是在降到大约6%时就停止了下降趋势，进入到了一个相对波动的阶段（见图3.10b）。如图3.10c所示，北京市的第三波疫情的这一平台传染速率（6%）比前两波疫情的平台传染速率0.4%足足大了1个数量级，这是与奥密克戎的超级传染率有关。这些特点说明，当前北京市的第三波疫情仍然处于较高位的传播阶段，对未来疫情演化的预测，提出了严峻的挑战。

但是，基于对国家干预力度演化规律的把握，GHR模型能够给出较为合理的疫情演化的预测。首先，根据当前疫情数据（截止于5月23日）的模拟结果显示，本轮疫情的隔离度于5月12日附近就已经接近饱和值96.4%了（见图3.10.a），与前两轮疫情的隔离度都迅速升至100%相比，这反映了北京市采取了“动态清零”政策在隔离度上包含一定的弹性和灵活性的特点。例如，在保障社会正常运转的情况下，实施了连续地大规模区域核酸检测，精准划分隔离区、封控区和防范区等综合隔离措施。根据GHR模型预测，新增确诊人数将于6月15日附近降至个位数，即疫情基本被控制，而累计确诊病例约为1964人（见图3.10d）。而如果此时考虑到适度放松隔离度，使饱和隔离度 $g_0 = 93\%$ ，这时疫情基本被控制的时间将延迟到7月6日，而代价是累计确诊病例也将增加804人。而如果此时进一步提高隔离度，使饱和隔离度 $g_0 = 98\%$ ，这会使疫情基本被控制时间提前到6月6日，累计确诊病例将减少至1572人。由此可以看出，采取不同力度的隔离措施将在一定程度上影响疫情走向，而隔离度的饱和值 $g_0$ 是定量刻画当前疫情隔离度的重要参数。

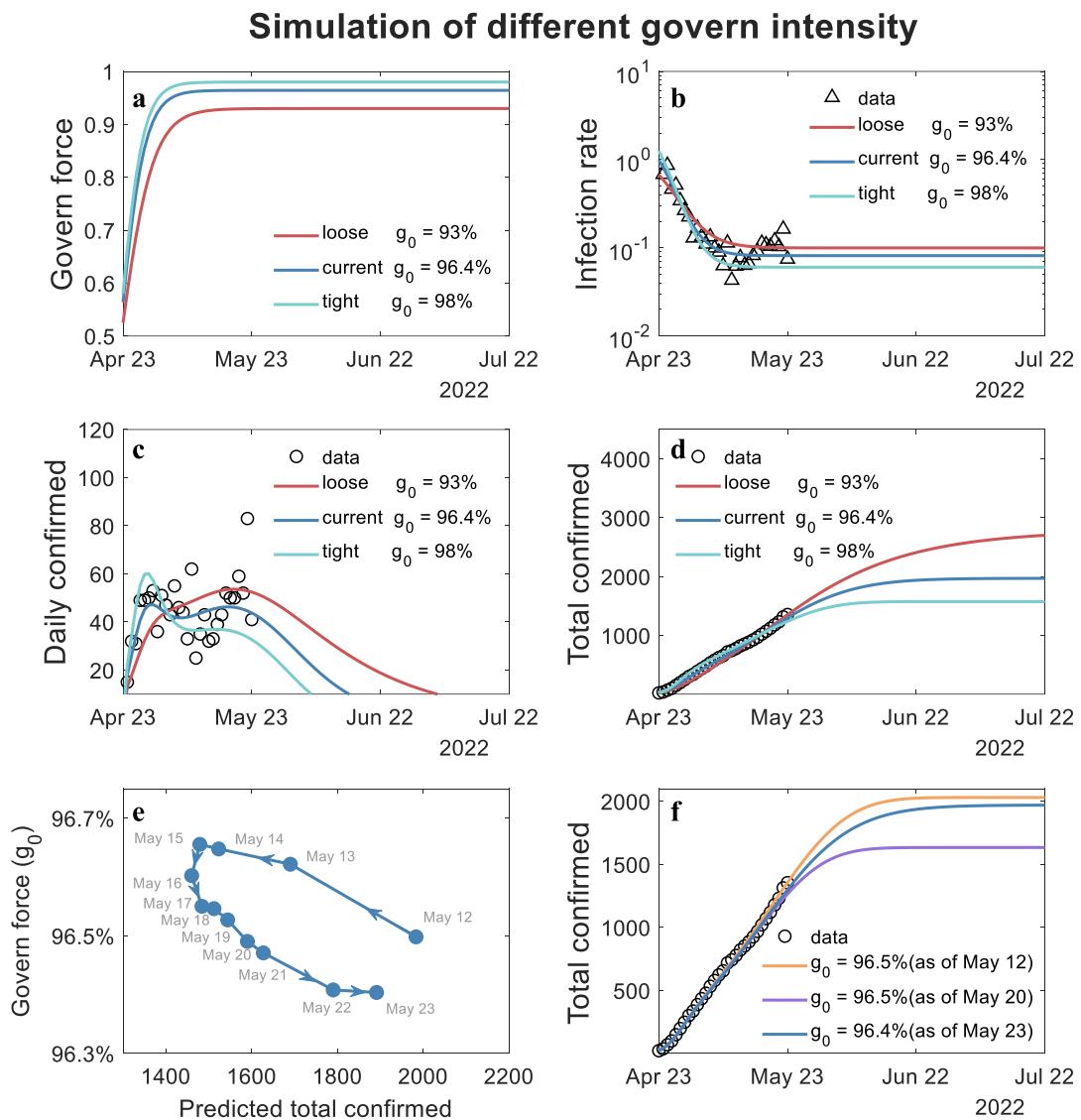


图 3.10 北京市第三波疫情在不同干预力度下的演化特征比较。黑色圆圈是来自于约翰·霍普金斯大学疫情实时监测系统收集的数据（截止于 2022 年 5 月 30 日），三角形是由此计算出的速率数据，实线则是 GHR 模型模拟结果，不同的颜色代表不同的饱和隔离度。a-b 分别为隔离度和累计确诊病例的演化结果。e 为采用不同大小的数据集预测的隔离度和感染人数。f 分别为使用截止于 5 月 12 日、20 日和 23 日的数据集的预测结果。

为了验证模拟预测的可靠性，我们选取了截止于 5 月 12 日-23 日之间不同日期的疫情数据，形成了 12 个不同大小的数据集，并运用 GHR 模型进行了预测（见图 3.10e, f）。如表 3.5 所示，在连续 12 天的预测中，模型参数表现得十分稳健，如关键的饱和隔离度  $g_0$  仅在 96.4%-96.7% 之间波动。值得指出的是，本模型是根据疫情发展数据反推隔离度的饱和值参数，目前还不具备条件从社会隔离政策直接估算该模型参数。因此，此模型还只具备描述性意义。

表 3.5 基于不同大小的数据集得到的模型参数的平均值

$\gamma$	$g_0$	$G(1)$	$\alpha$	$H(1)$	Confirmed
$0.33 \pm 0.01$	$96.5\% \pm 0.1\%$	$0.562 \pm 0.001$	$0.071 \pm 0.005$	$0.018 \pm 0.001$	$1674 \pm 172$

更重要的是，GHR 模型生动地展示了北京市以“动态清零”为目标而进行调整隔离度的过程。如图 3.10.e 所示，其中横坐标是使用不同数据集预测的最终累计感染病例数，纵坐标是对应的隔离度的饱和值  $g_0$ 。可以看到，5 月 12 日的隔离度为 96.5%，处于一个相对较低的值，预测的累计确诊病例数为 2028 人。这一预测结果，与截止于 5 月 23 日的预测结果十分接近，并且十分精确地刻画了这之后连续 12 天的累计病例的演化行为，平均预测相对误差只有 2.7%（见图 3.10.f）。但是，只有通过 GHR 模型展示的隔离度的日变轨迹，才能明确这背后是北京市针对隔离措施进行灵活调整的结果（见图 3.10.e）。例如，在 5 月 12 日后，北京市的隔离度逐渐连续增大，并于 5 月 15 日达到了隔离度的最大值 96.7%。由于确诊病例被发现存在一定的潜伏期，这一隔离度的增强实际反应的是更早一段时间的隔离措施的效果。根据媒体报道，北京市在 5 月初开展了多轮大范围的核酸检测，由此有效地提高了隔离度。而在此之后的 8 天内，北京市的隔离度开始逐渐下降，最初两天的下降速度很快，后来逐渐放慢了下降速度，并最终稳定在了在 96.4% 附近。这一隔离度的降低，得到了官方通报的多起聚集性病例的事实的支撑。经过这样一轮隔离度的提高和放松后，截止于 5 月 23 日的预测结果与 5 月 12 日的结果基本一致。考虑到北京市在此之后又提高了隔离度，如倡导更大范围的居家办公，因此未来实际的隔离度还会在此基础上有一定程度的增加，这样实际确诊病例会比当前的预测值稍低一些。例如，隔离度增长到 96.5% 时，最终的演化结果就会更接近于基于 5 月 20 日做出的预测。基于上述分析，我们认为当前的预测结果是可靠的。

由此，我们可以看出，目前北京市采用的防控策略是根据“动态清零”为目标导向，灵活地动态调整隔离度，从而在尽量减少对社会正常运转的影响下，始终将疫情控制在一定范围内，这为千万级人口大城市的可持续抗疫提供了宝贵的经验。而 GHR 模型通过对国家干预力度开展建模分析，为“动态清零”政策的有效实施，提供了定量评估当前干预效果和未来演化方向的工具。

### 3.5 本章小结

新冠疫情（COVID-19）是全世界瞩目的重大事件。新冠疫情传播演化规律的建模

研究，有助于人们开展有效的抗疫措施，减少感染和死亡。同时，新冠疫情本身是一个复杂系统，它为复杂性研究提供了丰富的研究数据和不同的演化场景。本章将人体序参量理论应用于群体的疫情演化行为。通过开展对称性分析，发现新冠疫情的传播演化中具有多层次的对称破缺的过程，即疫情早期由于新病毒出现的人群对称破缺，以及疫情爆发后由于国家干预导致的三类关键人群的对称破缺。

在疫情早期，由于新病毒的出现，使得所有人群发生了对称破缺。传统的 SEIR 传染病动力学模型通过抓住了这一过程中的统计结构，即不同人群之间是以一个恒定的速率在转换，由此建立了微分方程系统。本文从人体序参量理论的角度，更加深入地理解了 SEIR 传染病动力学模型的有效性的原因。

这次疫情爆发有二大特征，一是国家社会的有力干预，使得三类关键人群再次发生了对称破缺；二是公布的数据非常详实，便于开展定量研究。本文对社会干预力度进行了定量刻画，定义了（对广大人群进行隔离的）隔离度（ $G$ ），（对轻症患者的）治愈度（ $H$ ）和（对重症患者的）抢救度（ $R$ ）三个序参量，对应于三类干预措施。数据表明，这三个序参量在疫情发展过程中大幅增强，使得相应的三类转换速率演化发生了数量级的变化（从而使传统模型失效）。本文正是通过抓住这三个序参量的 Logistic 演化行为，构建了更符合真实疫情演化情况的 GHR 传染病动力学模型。

本文通过对不同特征地区、不同流行波次和不同干预力度等重要的疫情演化场景开展精细地模拟和分析，揭示了新冠疫情演化中存在的“序参量-转换速率-群体数量”三个层次之间相互耦合的疫情演化规律，证明了国家主导的三类干预措施在疫情演化中扮演的重要作用。这一结果，对于指导制定有效的抗疫措施，避免更多人的感染和死亡，具有重要意义。

值得指出的是，本研究遵循复杂系统研究的方法论，即“多层表述，逐级定量，多次迭代，逐步近似”<sup>[23,32]</sup>。本文所给出的 GHR 模型只是这一近似的一步，它只是抓住了主导项。后期我们将逐步引进各类次导项，如更丰富的人群接触、交通网络等数据，将 GHR 模型到更广泛和更真实的人群演化中，为当前抗击疫情提供有力的支撑。

## 第四章 应用于新冠死亡演化系统

在上一章，本文将人体序参量理论应用于群体的演化，建立了 GHR 模型，并且开展了国内外疫情发展过程的精细对比。本章将在此基础上，重点关注死亡人群的演化，来探索新冠疫情演化的一些普适规律。

在世界卫生组织（WHO）宣布 COVID-19 为大流行以来的两周年时，截止 2022 年 3 月 11 日，已有 184 个国家或地区报告了超过 600 万人死亡。据《柳叶刀》的一项调查研究<sup>[100]</sup>估计，截至 2021 年 10 月 31 日，约有 520 万儿童失去了父母或者照顾者。这是大流行给人类社会留下的永久伤痛。因此，迫切需要从第一波中吸取经验教训，如定量评估疫情的严重程度和澄清决定性的医疗和社会因素，以帮助政策制定者在下一波疫情中避免更多人的死亡。

### 4.1 新冠死亡率的复杂性

“死亡率”是一个人们非常关切的话题，也是大流行期间经常出现在媒体报告中的词汇。但同时，这也是被媒体滥用的词汇之一。例如，在 2020 年 8 月 3 日，美国总统特朗普接受 HBO 新闻网专访时，就“美国死亡率”这一问题与记者陷入了“争执”。原来，记者采用的死亡率的计算方法是死亡病例的数量除以当地的人口数量，而特朗普的计算方法是死亡病例的数量除以确诊病例的数量，来证明美国的死亡率较低。而在 4 个月前，他却是拿死亡病例的数量除以人口数量，来证明美国的死亡率较低。

由此可以看出，在疫情的不同时期，采取不同的“死亡率”的计算方法，能够得出完全相反的结论。因此，有必要对“死亡率”这个概念进行厘清和阐释。

#### 4.1.1 三种死亡率及其影响因素

实际上，所谓“死亡率”是一个很模糊的概念，真正需要明确是，究竟是以什么方式计算出的死亡率。通常主要有以下三种计算方法：

##### 病死率（Case Fatality Rate）

病死率的计算公式如下：

$$\text{病死率} = \frac{\text{死亡病例数}}{\text{确诊病例数}} \times 100\% \quad (4.1)$$

通过上式可以看出，病死率很受确诊病例数的影响。在疫情初期，由于许多危重症

患者被纳入医疗系统，而大量的无症状感染者和轻症患者还没有被确诊，所以使得病死率较高。而随着疫情的发展，开展了大规模检测筛查，使得确诊病例数增长很快，就会使得“病死率”曲线被降低。例如，在疫情刚爆发不久的 2020 年 2 月底，世界卫生组织（WHO）粗略估计每 1000 名确诊的新冠肺炎患者中就有 38 人死亡。而在几个月后的多项研究结果更趋近于 0.6%-1%<sup>[101]</sup>。

为了获得更准确的病死率估计值，人们在特定范围的人群内开展全员“血清阳性率”的调查，以此来消除“无症状感染者”对病死率估计的影响。多项研究结果估计病死率在 0.5%-1% 之间<sup>[102,103]</sup>。

### **粗死亡率（Crude Death Rate）**

粗死亡率的计算公式如下：

$$\text{粗死亡率} = \frac{\text{死亡病例数}}{\text{每 10 万人口数}} \times 100\% \quad (4.2)$$

由于在疫情期间，当地每 10 万人口数的变化通常不会很大，所以粗死亡率的变化主要是由死亡病例数的变化所引起的。因此，相对病死率，粗死亡率反映的主要是来自死亡病例的特点。本章也采用了“粗死亡率”来开展建模分析。

### **超额死亡率（Excess Mortality Rate）**

超额死亡率的计算公式如下：

$$\text{超额死亡率} = \frac{\text{实际死亡人数} - \text{预期死亡人数}}{\text{每 10 万人口数}} \times 100\% \quad (4.3)$$

这里的实际死亡人数是地区内包括与疫情相关或无关的所有死亡人数，预期死亡人数是根据往年没有发生疫情时估计该地区的平均死亡人数。二者的差值，就可以估计出由于疫情因素造成的额外死亡人数。这是一个在紧急时期能够很快统计出事件严重程度的量。例如，《自然》在 2020 年 9 月做出的分析表明，不同国家的超额死亡人数存在巨大差异。美国和西班牙分别有 25% 和 35% 的超额死亡人数，而秘鲁高达 74%。但同时，还有一些国家，比如保加利亚的超额死亡人数却是负值<sup>[104]</sup>。

这反映出超额死亡率在衡量疫情严重程度时，仍存在一些需要考虑的复杂性。首先，超额死亡率并没有区分因新冠病毒造成的死亡与因疫情间接因素造成的死亡。例如，美国疾病控制和预防中心（CDC）的一份报告<sup>[105]</sup>显示，在疫情早期，美国急诊量下降了 40%。但同时，糖尿病死亡人数比 5 年平均水平高出 20%-45%，死于缺血性心

脏病的人数比正常水平高出 6% 到 29%。这说明严重的医疗系统负担，使得许多非新冠患者由于得不到及时救治而死亡。

其次，由于严格的封锁和卫生习惯的改变，疫情同时还会造成死亡人数在其他方面的降低。例如，全球流感监测系统 FluNet 显示，2020 年的流感季缩短了一个多月。另外，疫情期间的交通事故、暴力犯罪等也有降低。这就一定程度上解释了为什么保加利亚的超额死亡人数是负值<sup>[104]</sup>。

总之，超额死亡率对于快速了解疫情严重程度的初步情况是很有帮助的，但其中存在着诸多复杂因素，这需要更丰富的数据支持和分析，才能较为准确的确定出超额死亡率。

#### 4.1.2 关键问题：各国死亡率的巨大差异

J 基于可靠的死亡统计数据开展定量建模，对于了解大流行是至关重要的<sup>[104]</sup>。然而，各国死亡率的演化呈现出巨大的差异。根据约翰·霍普金斯大学的全球疫情监测系统显示，无论是病死率还是粗死亡率，各国之间的差异达到了 2-3 个数量级<sup>[106]</sup>。

为什么各国的死亡率有如此大的差异？这是一个非常重要的问题。这一巨大差异迫使我们重新思考各国应对公共卫生危机的方式，并从中吸取教训，以提高全球应对这种规模的健康危机的能力。

然而，目前的研究还没有就这个问题得出一致的结论。具体来说，患者层面的研究表明，老龄、男性和有基础疾病的患者有更大的死亡风险，需要在重症监护室（ICU）进行呼吸道援助<sup>[107-112]</sup>。然而，确诊病例的人口统计特征一直随着时间的推移而发生变化。例如，与疫情爆发的早期阶段相比，后期阶段有更多的年轻人被感染。

另一方面，国家层面的研究表明，医疗资源可用性、感染规模等与死亡率有关<sup>[113-116]</sup>。然而，各国的预防、检测和应对疫情的能力在不同时期有很大的差异<sup>[117]</sup>。因此，COVID-19 的死亡率在空间（不同地区之间）和时间（不同疫情阶段）上呈现出极大的异质性<sup>[118]</sup>。所以，常用的横断面的研究方法，即选择处于不同演化阶段的某一天的死亡率数据进行横向比较分析，会产生误导性的表述。例如，两项发表的研究对病死率与检测次数之间是否存在关联，得出了不同的结论<sup>[113,114]</sup>。

此外，这些在不同层面分析的各种因素必须形成一个统一的观点，以便让人们更加清晰死亡率的演化图像，从而为抗击大流行提供有效的指导。

#### 4.1.3 已有研究需要实现的突破

为什么很多研究明确地提出了“横断面研究”这个方法的局限性，却又依然选择这一研究方法呢？

本文认为，一方面，这反映了死亡率本身的复杂性。如图 4.1 显示，各国的粗死亡

率具有十分复杂的动力学演化行为。例如，目前处于病死率第一的巴西，在疫情爆发后的很长一段时间内，相对其他国家的病死率都是比较低的。并且，它也没有明显的 Logistic 的“平台期”。这些特点，反映了各国在“病毒-个体-国家”多个层次因素耦合作用下的差异。因此，目前的研究也只能通过“横断面”的研究方法，来进行初步的统计分析研究，而难以开展动力学研究。

另一方面，这也反映了流行病学家们的一个根深蒂固的认识，即认为存在一个所谓的“真实”的死亡率，并基于它来反映新冠病毒的严重程度<sup>[101]</sup>。例如，上节提到的对三类死亡率的研究，都是在致力于确定出一个死亡率的“常数”。这一认识，跟上一章所介绍的认为转换速率是“常数”的认识是一脉相承的。那就是，因为疫情是由于病毒所引起的，所以人们更关注反映“病毒”本身特征的性质。

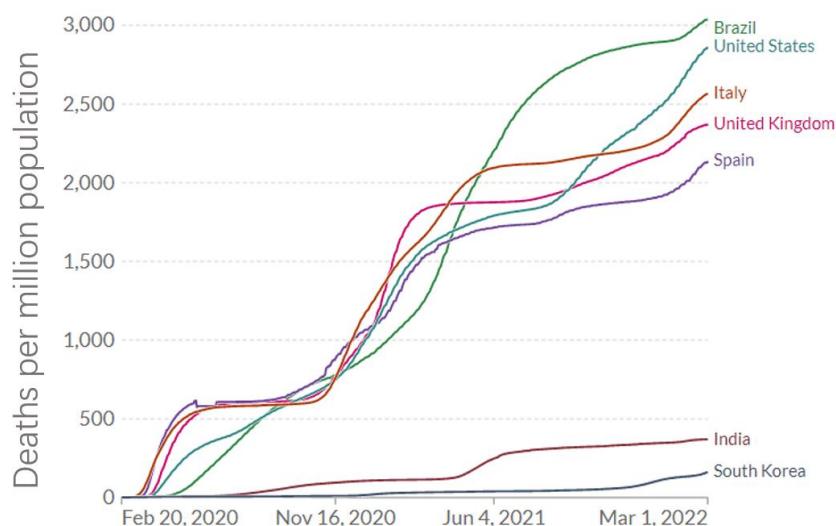


图 4.1 各国新冠粗死亡率的演化曲线

从“一元二面”的系统论视角来看，新冠病毒的致死力只是“群体死亡”这个系统中的一面，另外一面是外界施加的干预。随着科学技术的发展，这一面在现代国家的抗击疫情中扮演着越来越重要的角色。

因此，我们需要更新原来的认识，在同时把握这两面的基础上，来重新认识“群体死亡”这个系统。所以，本章采用了动力学研究，对全球第一波死亡率的全周期演变进行了完整的刻画，以更好地发现其中存在的主导序参量并解释其内在机制，从而发现对大流行演变中的普适规律。

## 4.2 新冠死亡率的序参量模型

如上所述，新冠死亡率的演化充满了复杂性。在这里，我们采取了第一章所介绍的

“从定性到定量综合集成”的方法来开展研究。具体来说，首先开展定性分类，区分具有不同动力学演化特征的国家，然后对同一类国家死亡率的演化行为进行定量刻画，并从中提取出主导序参量，从而开展不同序参量的关联分析，发现死亡率不同演化特征的内在机制。

#### 4.2.1 定性分类不同动力学特征

首先，由于政府政策变化、对数据的多次修正等因素，一些国家的死亡率演变显示出复杂的动力学演化行为（见图 4.1）。例如，有的国家在第一波疫情结束之前就爆发了第二波疫情，而有的国家在第一波疫情基本结束后的很长时间之后，才爆发了第二波疫情。这反映了背后不同的动力学演化因素。因此，首先必须对这些国家根据其动力学特征开展分类。

其次，随着疫情的发展，病毒变异对于死亡率演化具有不可忽视的影响，所以我们选择只分析具有完整的第一波死亡率演化行为的国家。为此，我们需要明确第一波疫情的持续时间段。同时，为了防止在比较和分析各国大流行的演变时间时出现长尾效应，本文选择了新增死亡病例的峰值前的  $1/10$  峰值处所对应的时间为疫情的爆发日期 ( $t_{outbreak}$ )，峰值后的  $1/10$  峰值处对应的时间为疫情的饱和日期 ( $t_{saturation}$ )，如图 4.2a 所示。爆发日期与饱和日期之间的时间，为第一波疫情的持续时间。由于每日新增死亡人数的涨落较大，所以对其进行了 10 天的平滑处理。

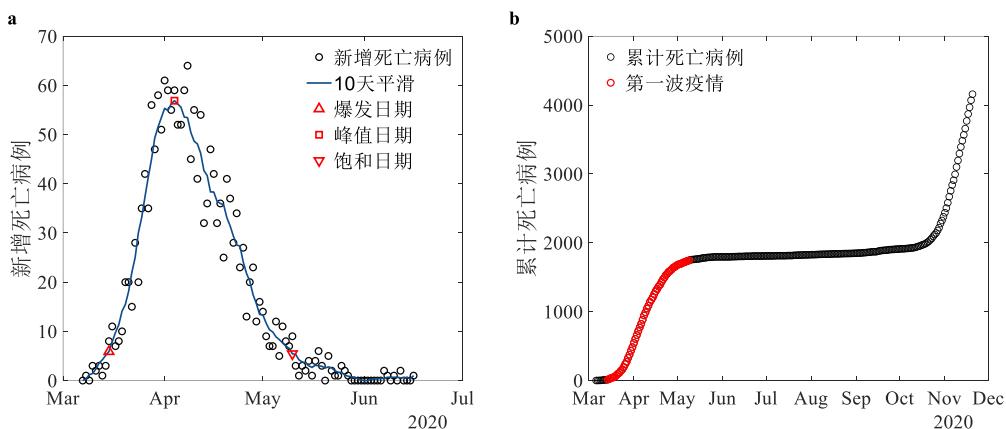


图 4.2 第一波疫情的死亡率的完整演化周期（以瑞士为例）。a 每日新增死亡病例数和特征日期。黑色圆圈是报告的每日新增死亡病例数。蓝线是每日死亡人数的 10 天平滑线。正方形是每日新增死亡病例数的峰值。正三角形是新死亡病例数攀升至峰值的十分之一以上时的爆发日期。倒三角形是新增死亡病例数下降到峰值的十分之一以下时的饱和日期。b. 黑色圆圈是报告的累计死亡人数。红色圆圈是第一波疫情的数据。

最后，考虑到统计学的稳定性，合格的样本应该有超过 100 例的累计死亡人数。由此，本文提出了明确的纳入标准，即具有完整的第一波死亡率演化周期并且第一波

的累计死亡病例数大于 100 例的国家或地区，才被视为合格的纳入样本。最终，共计 52 个国家或地区被纳入本项研究，如图 4.3 所示。

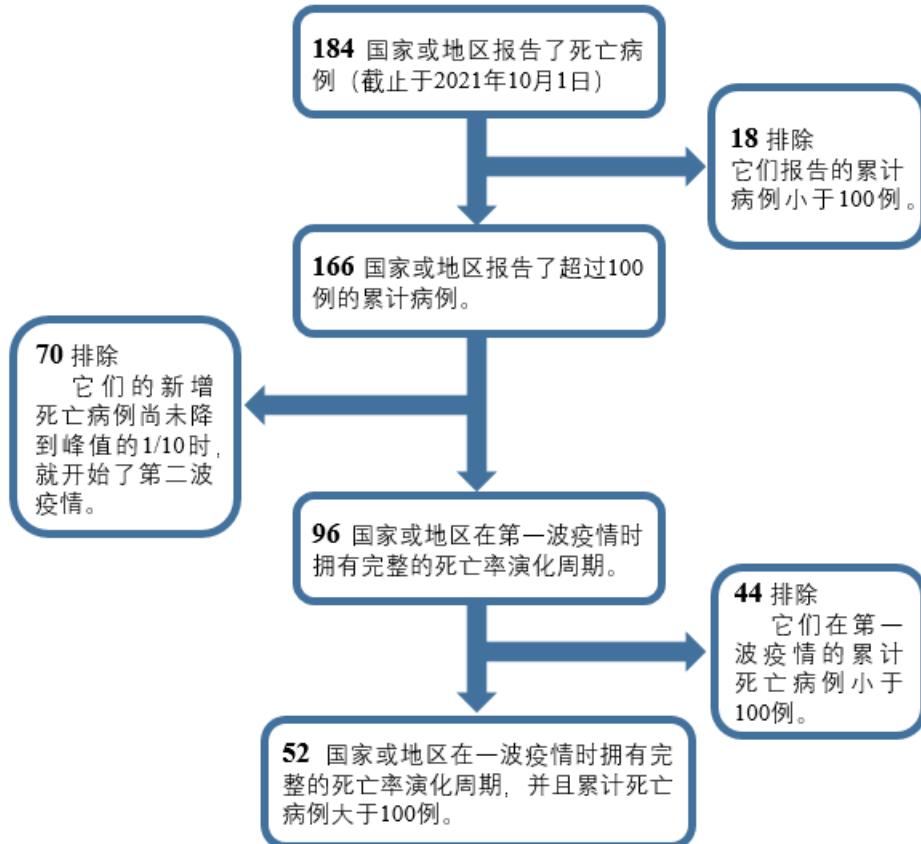


图 4.3 本章的纳入标准。

#### 4.2.2 定量刻画全周期演化规律

由图 4.1 可以看出，对于在第一波疫情具有完整的死亡率演化周期的国家，其死亡率的演化行为符合 Logistic 演化形式。这是因为，在疫情爆发的早期阶段，由于干预不力，死亡人数几乎呈指数级增长。随着死亡人数的扩大，社会系统采取行动减缓了死亡率的增长，如增加可用的医疗资源，控制疫情的蔓延，保护高危人群。最终，死亡人数在大流行的后期趋于饱和。

这一结论，也可以从 GHR 模型中推导出来。式 3.15 可以写成：

$$\Delta D(t) = S(t) * Q(t) \quad (4.4)$$

其中， $\Delta D(t)$  是新增死亡病例数， $S(t)$  是死亡速率， $Q(t)$  是现有确诊病例数。

因此，上式两边取积分就是累计死亡人数  $D(t)$ ，再除以人口数  $P$ ，就可以得到粗死亡率  $M(t)$ ：

$$M(t) = \frac{1}{P} \int S(t) * Q(t) \quad (4.5)$$

由式 3.14 可以得到，在抢救度  $K(t)$  达到饱和值  $k_0$  后，死亡速率  $S(t)$  则为一个常数。这一点在国内外的疫情数据（见图 3.6-3.10）中都得到了验证：

$$S(t) = s_0(1 - k_0) \quad (4.6)$$

其中， $s_0$  是初始死亡速率。

因此，4.5 式可以写成：

$$M(t) = \frac{s_0(1 - k_0)}{P} \int Q(t) \quad (4.7)$$

上式中， $s_0(1 - k_0)/P$  是一个常数，而积分量  $\int Q(t)$  满足 Logistic 演化模式。所以，粗死亡率  $M(t)$  也满足 Logistic 演化模式。

因此，我们使用 Logistic 模型来拟合 COVID-19 的粗死亡率（每 10 万人口的死亡人数， $M(t)$ ）：

$$\frac{dM(t)}{dt} = \frac{1}{\tau} M(t) \left(1 - \frac{M(t)}{m_0}\right) \quad (4.8)$$

方程 (4.8) 的解可以写成如下形式：

$$M(t) = \frac{m_0}{1 + e^{(t_m-t)/\tau}} \quad (4.9)$$

#### 4.2.3 模型参数的意义

##### 饱和死亡率 $m_0$

在死亡率接近饱和时，即当  $t \gg t_m > 0$  时，4.9 式可以写成：

$$M(t) \approx m_0 \quad (4.10)$$

因此， $m_0$  表示死亡率演化后期的饱和死亡率。

##### 饱和死亡率 $t_c$

当  $t = t_m$  时，4.9 式可以写成：

$$M(t_m) \approx \frac{m_0}{2} \quad (4.11)$$

因此， $t_m$  表示死亡率达到饱和死亡率的一半时的特征时间。

## 早期增长速率 $\tau^{-1}$

在疫情早期，即当 $t \approx 0$ 时， $M(t) \ll m_0$ ，方程（4.8）可写为：

$$\tau^{-1} \approx \frac{M(t)'}{M(t)} \quad (4.12)$$

这是一个指数增长函数，所以 $\tau^{-1}$ 表示死亡率演化早期时的指数增长速率。

综上，我们推导了刻画全周期粗死亡率的 Logistic 模型，并且解释了模型参数的物理意义。

特别需要注意的是，模型中只有两个参数，即表示早期死亡率演化的增长速率 $\tau^{-1}$ 和表示晚期死亡率演化的饱和死亡率 $m_0$ 。这样，就把复杂的死亡率演化行为，解耦为两个互相独立的阶段。

因此，虽然各个国家的死亡率演化行为各不相同，但是 Logistic 模型只用两个简单的参数，就能实现对它们的全部刻画。换句话说，各国死亡率演化的规律，都包含在这两个参数之中。而且，由于这两个参数互相独立，我们就可以分别研究每个参数的演化规律。

至此，我们通过运用“从定性到定量”的复杂系统的研究方法，就使得复杂的死亡率演化难题被解耦了。

### 4.2.4 刻画 52 个国家的死亡率数据

在使用模型刻画数据、开展分析之前，这里首先对本章所使用的数据，进行一个简要的介绍。所有数据都来自开放的数据库，具体来说：

COVID-19 病例数据是由约翰-霍普金斯大学的数据库汇总而来<sup>[106]</sup>（截至 2021 年 10 月 1 日）。而武汉市的数据来自于湖北省卫生与健康委员会收集的数据<sup>[119]</sup>。

各国的人口、年龄中位数和医院床位数（每千人）来自于著名的公共数据库 Our World in Data<sup>[120]</sup>。对于中国的数据，我们选择了武汉市的数据来代替，而不是全国的数据。因为武汉市是疫情爆发中心，并且它的人口规模、大流行规模与其他国家或地区相当。武汉市的人口和床位数据来自武汉市统计局<sup>[121]</sup>。值得一提的是，这些国家级变量的数据是更新到最近一年的，但往往不是同一年的。

我们用 Logistic 模型（4.9 式）拟合了 COVID-19 第一波疫情中的 52 个国家的全周期死亡率数据，如图 4.4 所示。由图可以看出，52 个国家的死亡率演化行为，都能被 Logistic 模型很好地刻画，证明了模型的有效性。

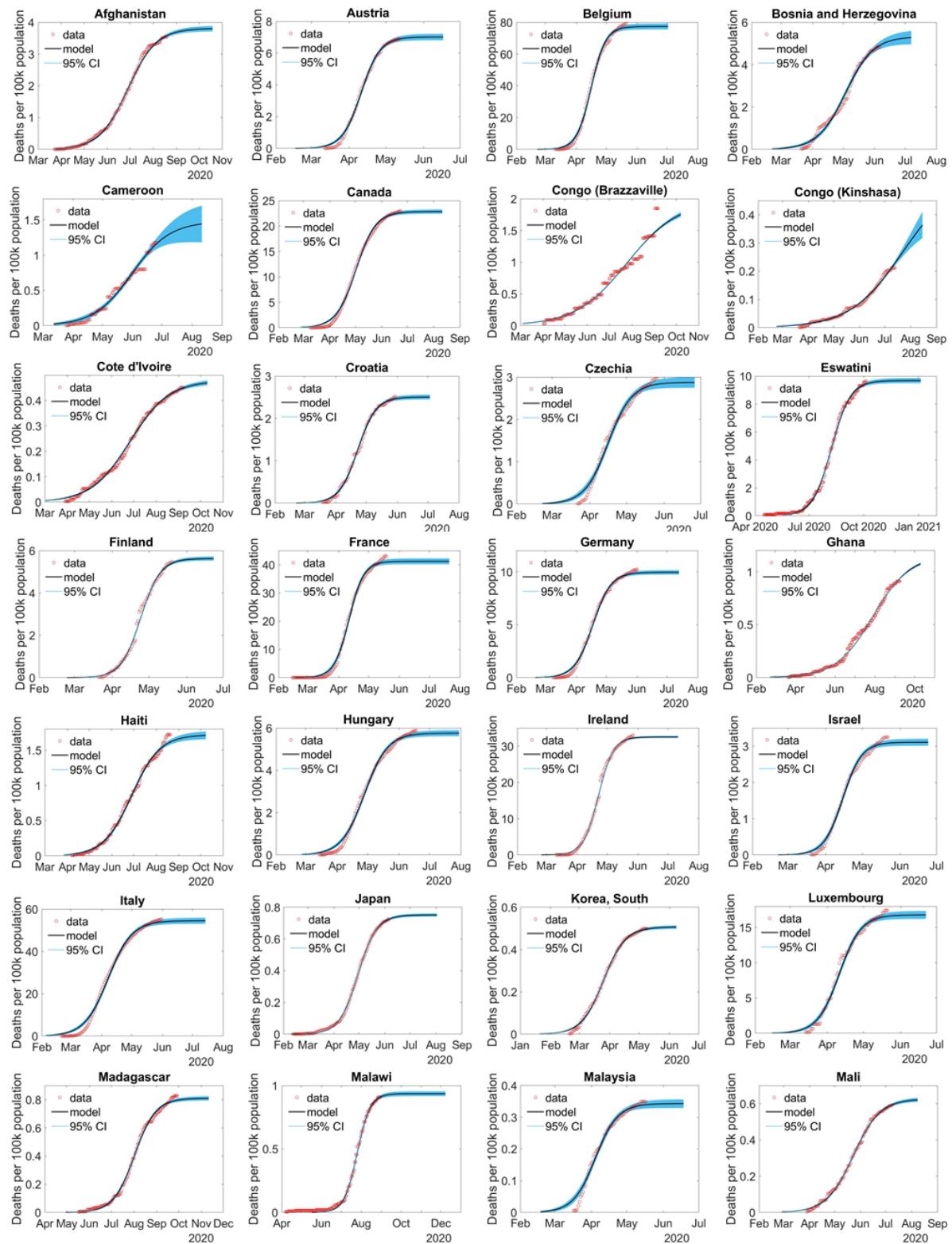


图 4.4 模型对各国第一波全周期死亡率的拟合结果。红点是官方报告的数据，黑线是有 95% 置信区间（蓝带）的模型拟合结果。（转下页）

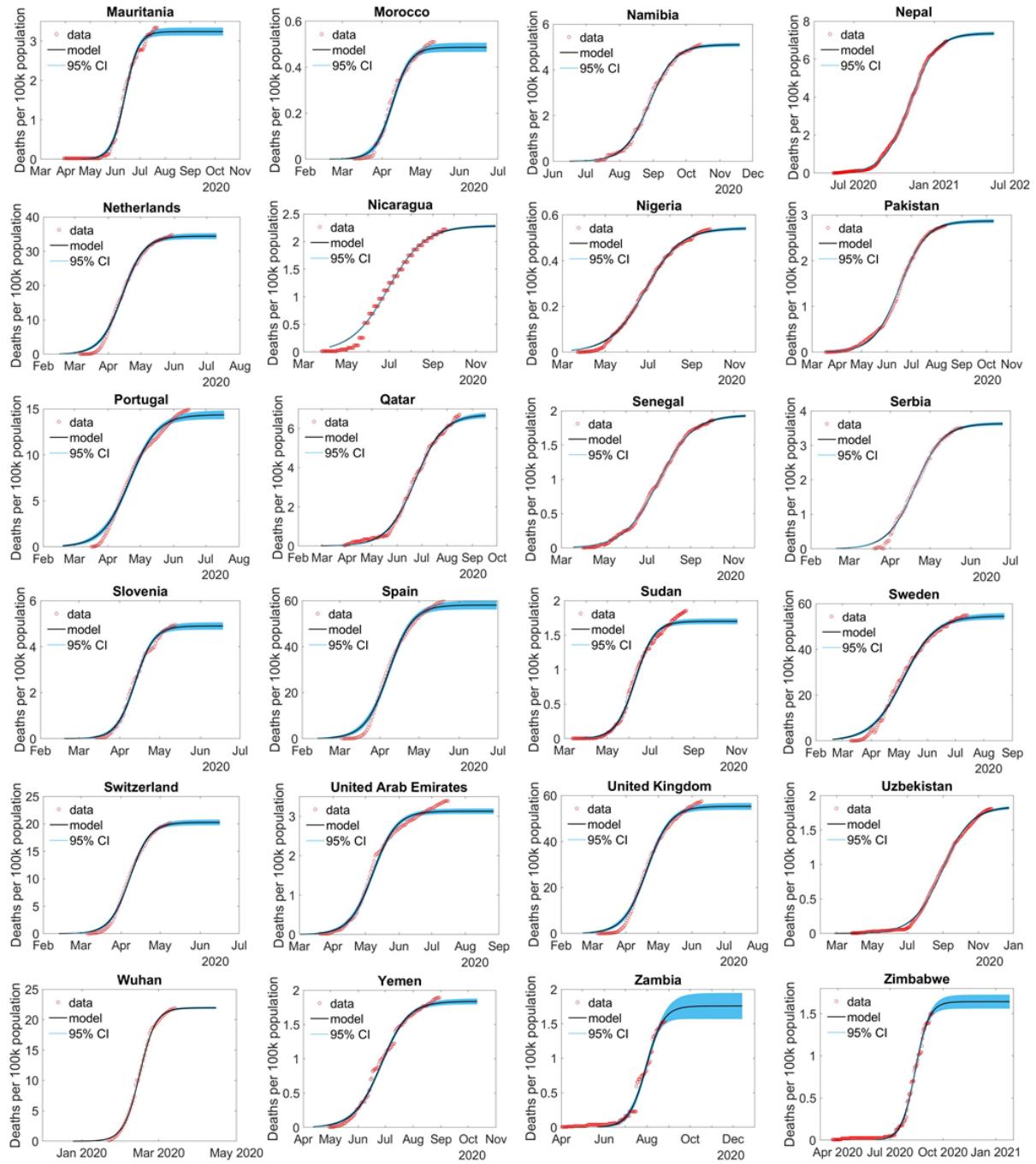


图 4.4 模型对各国第一波全周期死亡率的拟合结果。红点是官方报告的数据，黑线是有 95% 置信区间（蓝带）的模型拟合结果。（接上页）

### 4.3 饱和死亡率的跨尺度特征

#### 4.3.1 与年龄、床位占有率的幂次律

当疫情结束时，死亡率  $M(t)$  演变为  $m_0$ ，所以饱和死亡率参数  $m_0$  主要量化了各国疫

情死亡人数的特征。根据已经发表的临床研究<sup>[108-112]</sup>，年龄是影响新冠死亡风险的重要因素。所以，我们在图 4.5a 中绘制了  $m_0$  作为国家年龄中位数 (A) 的函数，它立即显示出一些显著的特征：高死亡率模式 ( $m_0 > \bar{m}_0 = 12.0$ ) 都出现在年龄中位数超过 38 岁的国家，而年龄中位数小于 38 岁的国家都属于低死亡率模式 ( $m_0 < \bar{m}_0 = 12.0$ )。

这表明年龄中位数 (A) 是影响死亡率的关键因素之一。这与临床观察到的结果一致，而我们的研究则更加地量化了这一关系，即饱和死亡率  $m_0$  和年龄中位数 (A) 之间存在着幂次律：

$$m_0 = 0.003 * A^{2.1} \quad (4.13)$$

其中，A 指国家年龄的中位数。

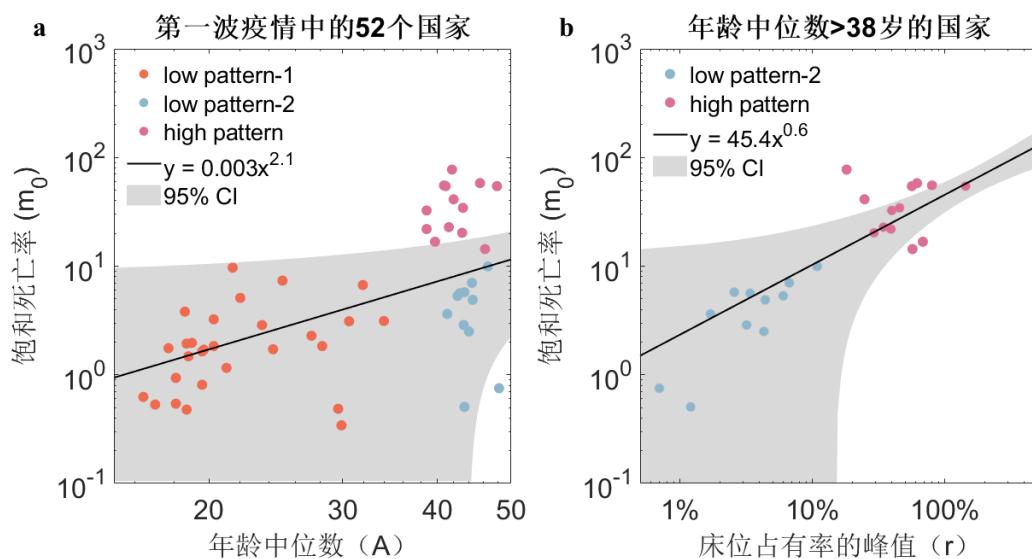


图 4.5 年龄中位数和床位占用率的峰值解释了死亡率的巨大变化。(a) 各国年龄中位数与饱和死亡率之间的幂次律 ( $N = 52$ )。(b) 年龄中位数在 38 岁以上的国家的饱和死亡率和床位占用率的峰值之间的幂次律 ( $N = 24$ )。橙色、蓝色和粉红色的点分别对应于低死亡率模式-1、低死亡率模式-2 和高死亡率模式。黑线是幂函数的拟合预测值。灰色区域为 95% 的置信区间。

(a) 和 (b) 都采用了双对数坐标，以更好地显示数据。

另一方面，我们发现对于年龄中位数高的国家来说，它们与这一幂律的偏差相当大，即图 4.5a 中的粉红色的点大都在置信区间之外。这种强烈的涨落表明，一定存在其他的重要因素导致了高死亡率。

进一步，我们发现对于中位数年龄超过 38 岁的国家，饱和死亡率  $m_0$  和床位占有率为峰值 ( $r$ ) 之间也存在一个很好的幂次律（如图 4.5b 所示）：

$$m_0 = 45.4 * r^{0.6} \quad (4.14)$$

其中， $r$  指床位占用率的峰值，根据在院确诊病例的峰值除以医院床位数（每千人）而

获得。

饱和死亡率 $m_0$ 随着床位占用率 $r$ 的峰值呈幂次律增加。这一关系将老龄国家死亡率的巨大变化，归因于床位占用率峰值的巨大变化（几乎为两个数量级）。根据床位占用率是否大于 15%，可以将老龄国家划分为两组，而大于 15% 的老龄国家都是高死亡率模式，小于 15% 的老龄国家都是低死亡率模式。

值得一提的是，4.5b 中离置信区间最远的红点代表的是比利时。它之所以离幂律曲线较远，很可能是因为比利时针对 COVID-19 的死亡病例，采用了更广泛的纳入标准<sup>[122]</sup>。

#### 4.3.2 三类典型的演化模式

在上述结果的基础上，各国死亡率的巨大差异，就可以得到定量的解释。首先，年龄中位数和床位占用率发现的两个关键阈值，可以将所有国家分为三个典型模式。图 4.6 显示了 52 个国家的死亡率模式是如何在不同大洲、年龄中位数和床位占用率之间演化的。橙色、蓝色和粉红色的连接线分别对应于低死亡率模式-1、低死亡率模式-2 和高死亡率模式（与图 4.5 中的颜色一致）。每条线的宽度与国家的数量成正比。

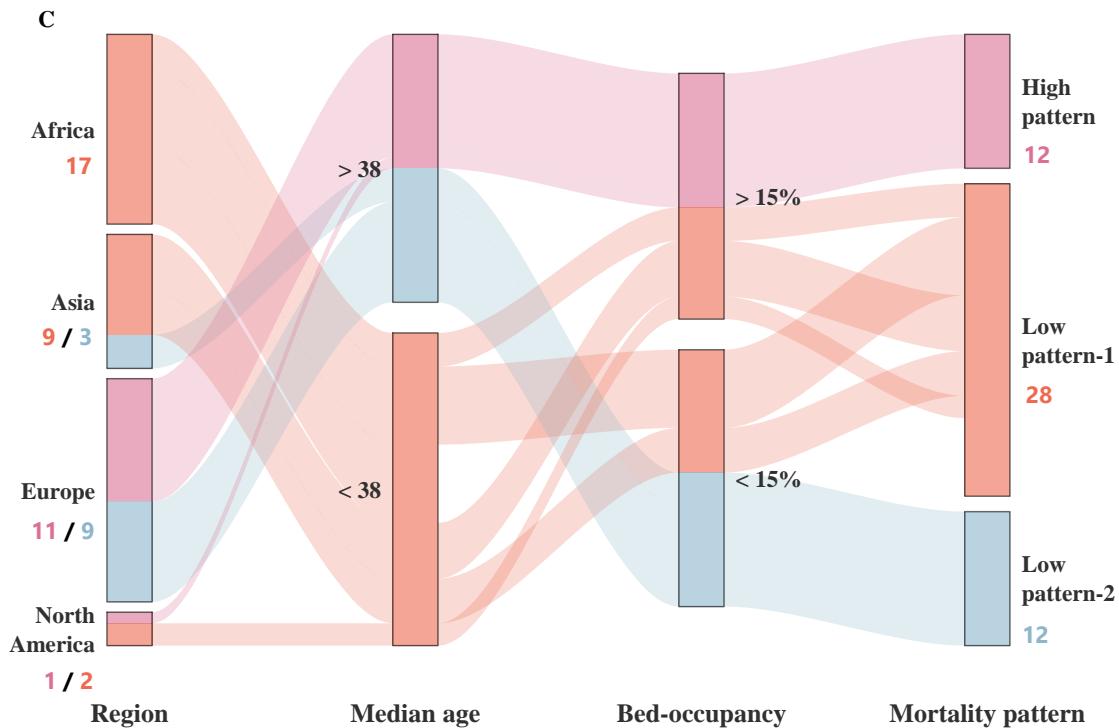


图 4.6 三类典型的死亡率演化模式。

首先，病床占用率高 ( $>15\%$ ) 的老龄国家 ( $>38$  岁) 的死亡率远远高于其他两个低死亡率组，其中包括那些年轻的国家 ( $<38$  岁)，以及床位占用率低 ( $<15\%$ ) 的国家。

非常有意义的是,对于一些老龄国家( $>38$ 岁),如果它们的床位占用率较低( $<15\%$ ),也能够演变成低死亡率的模式(图4.5中的蓝色点)。这表明“床位占用率”是评估国家抗击COVID-19措施有效性的一个重要指标。

需要说明的是,我们提出用“床位占用率”来分析饱和死亡率,而不是很多研究所采用的“床位数”。这是因为,“床位占用率”更能反映各国医疗资源的可用性。已经发表的横断面研究的确显示了年龄、医院床位和死亡率之间存在着相关性,但不清楚这些因素在不同的国家和爆发的不同阶段发挥了哪些作用。我们的研究则提出了两个关键阈值,对于各国了解和监测本国疫情发展的危险程度,制定适合本国国情的抗疫政策,有重要的意义。

#### 4.3.3 国家干预措施的定量评估

根据上述饱和死亡率 $m_0$ 与床位占用率 $r$ 之间的幂次律(见图4.5b),我们就可以来定量地评估国家干预措施的有效性。

首先,对于死亡率较低的国家——如奥地利( $m_0 = 7.02$ )和德国( $m_0 = 9.95$ ),以及死亡率高的国家——如加拿大( $m_0 = 22.85$ )和瑞典( $m_0 = 54.54$ ),饱和死亡率和床位占用率之间的幂次律,都能够对这些国家的情况进行较为精确的刻画。

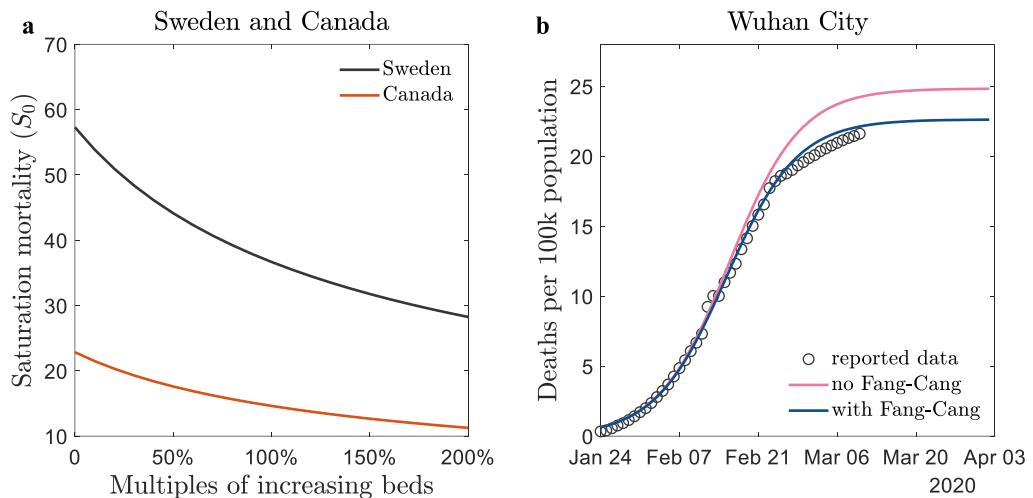


图4.7 死亡率随着床位数的增加的变化。**a** 改变床位数时饱和死亡率的降低。黑线和橙线分别代表瑞典和加拿大。**b** 武汉市(中国湖北省)修建方舱医院对降低死亡率的贡献。灰色圆圈是官方报告的数据。红线代表没有方舱医院的死亡率演变模拟,而蓝线代表有方舱医院的死亡率演变模拟。

有趣的是,根据这个幂次律可以预测,当病床数量增加一倍时,死亡人数将减少约36%(见图4.7a)。换句话说,增加病床的数量对于避免老龄国家更多的死亡病例是非常有效的。具体来说,如果加拿大可以将床位数量增加一倍,在第一波疫情中就可以避

免大约 3500 人的死亡。同样，瑞典可以避免大约 2200 人死亡。

另外，我们根据这一幂次律，定量地评估了武汉市修建方舱医院对于减少死亡病例的效果。这里选取武汉市（而不是湖北省），是因为它的人口规模和大流行规模与一些国家更加相当。

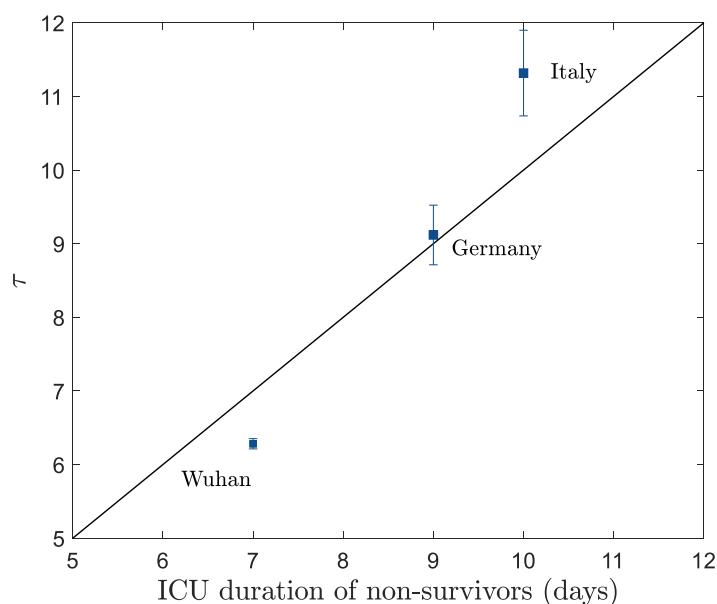
如果武汉市不修建方舱医院，根据爆发前的床位数计算的床位占用率的峰值为 39.2%，而据此根据式 4.14 预测的饱和死亡率 ( $m_0$ ) 为 24.8。然而实际上，根据媒体报道，武汉市通过修建方舱医院，增加了 15,000 张床位数<sup>[123]</sup>。根据这一新的床位数，通过式 4.14 预测的饱和死亡率  $m_0$  将会减少到 22.6，这与实际的饱和死亡率 22.0 非常接近，即根据这一幂次律可以解释武汉市死亡人数下降的 76.7% (248 人)。这一结果有力地证实了上述幂次律的有效性。

## 4.4 疫情持续时间的跨尺度特征

上节讨论了表示死亡率演化晚期的模型参数  $m_0$ ，回答了为什么各国死亡率存在巨大差异的难题。本节将继续讨论另一个模型参数  $\tau$ ，它反映了死亡率演化的早期特征。

### 4.4.1 ICU 时间 $\tau$

首先，根据 4.12 式， $\tau^{-1}$  表示了死亡率演化早期的指数增长速率。它的倒数  $\tau$ ，则反映了新增死亡病例的特征时间。由于几乎所有 COVID-19 的死亡报告都发生在 ICU，所以本文建议  $\tau$  表示疫情早期的非存活者在 ICU 的时间。



4.8 ICU 的时间和  $\tau$  的值对比。45 度的黑线表示观察和预测数据的相等。

由于疫情早期的非存活者在 ICU 的时间在临床中并不是一个非常重要的量，因此我们在已发表的临床研究中只找到了武汉、德国和意大利三个国家或地区的临床观察数据<sup>[124-126]</sup>。根据对比，发现模型拟合数据得到的参数  $\tau$  与实际观察发现的非存活者的 ICU 时间之间十分接近，如图 4.8 所示。因此，这初步对  $\tau$  的物理意义进行了验证。

#### 4.4.2 无量纲系数 $k$

当疫情刚开始爆发时，即  $t = t_{outbreak}$ ，根据模型（4.9 式）有：

$$t_{outbreak} = t_m - \ln(m_0/M(t_{outbreak}) - 1) * \tau \quad (4.15)$$

当疫情接近饱和时，即  $t = t_{saturation}$ ，根据模型（4.9 式）有：

$$t_{saturation} = t_m - \ln(m_0/M(t_{saturation}) - 1) * \tau \quad (4.16)$$

式 4.16 减去式 4.15，则可以得到疫情的持续时间  $T_{duration}$ ：

$$T_{duration} = t_{saturation} - t_{outbreak} = k * \tau \quad (4.17)$$

上式中，

$$k = \ln\left(\frac{(m_0/S(t_{outbreak}) - 1)}{(m_0/S(t_{saturation}) - 1)}\right) \quad (4.18)$$

由此，我们可以看到，疫情持续时间  $T_{duration}$  与疫情早期非存活者的 ICU 时间  $\tau$  之间为严格的线性关系。而二者之间的联系为一个无量纲的系数  $k$ 。

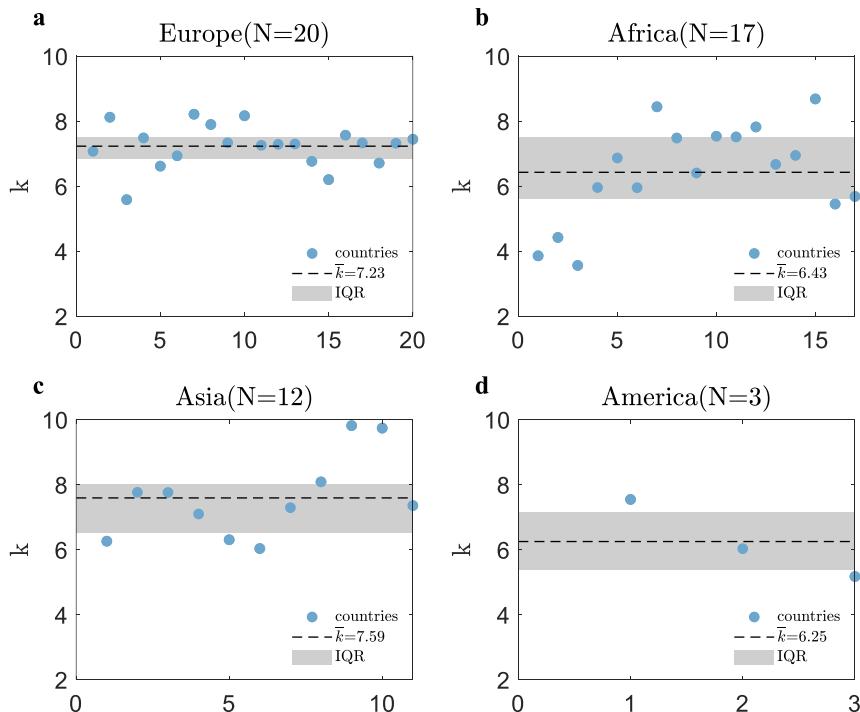


图 4.9 52 个国家的  $k$  值。a-d 分别是欧洲、非洲、亚洲和美洲国家的  $k$  值。蓝色圆点是各国的实际数据。黑色的虚线是各大洲国家的平均数。灰色区域是四分位数范围（第 25-75 位）。

我们计算了 52 个国家的  $k$  值，如图 4.9 所示。一个非常重要的发现是，除了个别国家，大部分国家的  $k$  值十分接近，其平均值为 7.0 (IQR, 6.3-7.7)。这里的一个主要原因是 4.18 式进行了对数运算。所以，虽然各个国家的饱和死亡率具有数量级的差异，但是经过对数运算后，这一差异被很大程度上“消除”了。因此，我们可以用一个“平均”的  $k$  值来刻画各个国家之间的线性关系。

#### 4.4.3 普适的线性率

在上述讨论的基础上，我们在图 4.10 中展示了各国家的 ICU 时间  $\tau$  和疫情持续时间之间的普适线性率，这为 COVID-19 的传播和演变提供了新的重要见解。

首先，这一线性规律在 ICU 时间较短的国家得到了更好的遵守 ( $\tau < 20$ )。而对于 ICU 时间较长 ( $\tau > 20$ ) 的国家，数据相对较为分散，这很可能是由于随着疫情时间的不断持续，这个过程中存在更多的影响因素。

这一结果表明，疫情早期的非存活者的 ICU 时间是反映疫情持续时间的一个很好的指标。疫情早期，人们对于新病毒的认识有限，所以非存活者的 ICU 时间主要反映了病毒本身的性质。而疫情的持续时间是一种社会规模属性。这两者之间存在一个普适的线性关系，这揭示了 COVID-19 的一个重要的性质，即存在时间的跨尺度特征。

如果病毒的毒性越强，则 ICU 病人很快就会死亡，即疫情早期的非存活者在 ICU 的时间就越短。而根据我们发现的线性律，这意味着整个疫情持续的时间也越短。相反，如果病毒的毒性越弱，整个疫情持续的时间则越长。

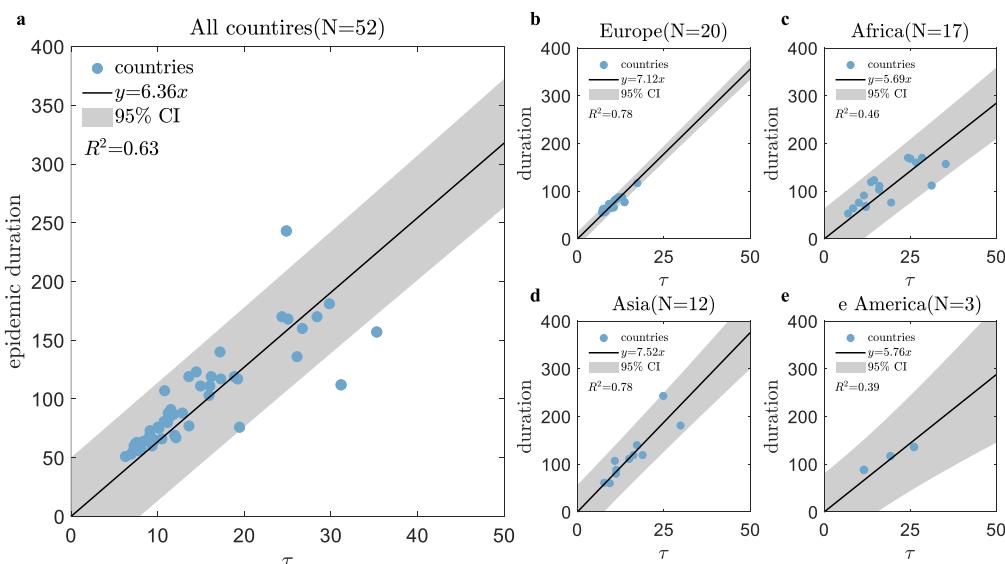


图 4.10 ICU 时间  $\tau$  与疫情持续时间之间的跨尺度关联。a-e 分别是所有国家、欧洲、非洲、亚

洲和美洲的线性率。横坐标是模型的拟合参数 $\tau$ ，表示疫情早期的非存活者的 ICU 时间。纵坐标变量是疫情的持续时间。蓝点是各国的数据。黑线为线性函数的拟合预测值。灰色区域为 95% 的置信区间。

#### 4.4.4 预测疫情持续时间

上述结果的一个直接应用，就是我们就可以在疫情早期根据普适的线性率来预测疫情持续的时间。如表 4.1 所示，我们使用无量纲系数 $k = 7 - 8$ 和文献<sup>[124-126]</sup>报告的 ICU 时间来预测疫情的持续时间。结果显示，普适线性率预测的疫情持续时间，与实际的持续时间非常接近。

表 4.1 基于普适线性率预测疫情持续时间 ( $k = 7 - 8$ )

国家/地区	ICU 时间（天）*	实际持续时间（天）	预测持续时间（天）
武汉	7	51	49-56
德国	9	73	63-72
意大利	10	87	70-80

\*数据以参考文献<sup>[124-126]</sup>中的中位数表示。

这一结果有力地证实了普适线性率的有效性。疫情的中长期预测一直是一个难题。我们的这一结果可以帮助政策制定者更准确地预估抗击大流行的困难，更好地优化卫生系统的负荷，以尽量减少死亡人数。

### 4.5 本章小结

本章通过将人体序参量理论应用于新冠死亡演化系统，定量地回答了人们在新冠大流行中关注的两个重要问题，即为什么各国的新冠死亡率有如此大的差异，以及当地疫情将持续多长时间？

对于第一个问题，本文通过发现的饱和死亡率 ( $m_0$ ) 与年龄中位数、床位占用率之间的幂次律（式 4.13 和式 4.14），对各国死亡率的巨大差异进行了定量解释（见图 4.5）。更重要的是，本文从中发现有两个关键的阈值——年龄中位数为 38 岁和床位占用率为 15%——将各国死亡率的演化划分为三种典型的模式（见图 4.6）。特别是，老龄国家可以通过将床位占用率降低到 15% 以下而演化为低死亡率模式。这一发现使我们能够产生一个具有实际意义的预测，若老龄国家（中位数年龄>38 岁）的病床数量增加一倍，可减少约 36% 的死亡人数（见图 4.7）。以上结果对于各国在下一波疫情中评估死亡风险，准备充足的医疗资源，提供了定量的支撑。

对于第二个问题，本文通过发现的两个不同时间尺度之间的线性律，使得在疫情

早期使用 ICU 数据来预测当地疫情的持续时间成为了可能（见图 4.10）。换句话说，通过疫情早期非存活者在 ICU 的时间，乘以大约 7 倍，就是当地疫情的持续时间。这一结果已经通过部分国家的临床数据进行了可靠的验证（见表 4.1）。众所周知，因为有许多复杂因素的存在，所以对疫情的中长期预测一直是一个巨大的挑战。事实上，我们也确实发现，一些疫情持续时间较长 ( $\tau > 20$ ) 的国家在线性律中的涨落较大。这很可能是因为这些国家在较长的疫情发展过程中，会经历更多的一些影响事件（见图 4.10）。但即便如此，全球 52 个国家仍然都符合一个几乎普适的线性律。这表明，时间跨尺度关联是 COVID-19 大流行的一个重要特征。未来一个很有意义的课题是，比较不同传染病——例如 SARS 或流感——的无量纲系数  $k$  的值，这可能为理解不同传染病演化规律带来新的启示。

本文的另一个重要发现是，早期非存活者在 ICU 的持续时间是一个十分重要的变量。但是，这一点在过去的临床研究中是经常被忽视的。例如，本文提出的模型所刻画的武汉市的非存活者的 ICU 时间在 52 个国家和地区中是最小的（6.3 天），这与临床研究中总结的 7 天非常接近<sup>[126]</sup>。这一结果与武汉市是疫情的第一个爆发点，当时人们对该病毒了解甚少的事实也是一致的。此外，各国的 ICU 时间之间存在较大的差异，这也需要在未来进行研究和解释。本文呼吁临床研究收集更准确的 ICU 时间数据，以便更好地预测下一波的大流行。

上述结果的获得，主要得益于人体序参量理论的应用。如前所述，已有的研究主要是通过横断面研究的方法，来寻找与死亡风险相关的因素。而本文是从动力学的视角出发，旨在对死亡率的全周期演化行为有一个完整的理解。首先，通过定性分类，挑选出了在第一波疫情中具有完整演化周期的 52 个国家。其次，根据序参量模型将死亡率的复杂演化行为解耦为疫情不同阶段的两个独立过程。这使得我们能够分别地研究疫情不同阶段的演化规律。未来，我们将继续对目前的模型进行测试，如果成功的话，将大大增强我们理解和控制当前大流行的能力。

上述研究还存在一些局限性。首先，这里分析的数据是来自公共数据库。数据库的质量有限、统计标准不同、不完整，可能会影响我们描述的精确性。第二，我们的研究不包括具有复杂演化行为的国家，如疫情的反复反弹。第三，我们使用了国家层面的年龄中位数数据，而不是与疫情更相关的病人层面的数据，这是因为目前大多数国家都缺乏此类数据。最后，我们使用了疫情爆发前的床位数，而在疫情爆发期间各国往往会上增加床位数——这些数据目前很难获得。尽管这些因素影响了描述的精确性，但它们不会影响对大多数国家第一波疫情的研究结论。

总之，我们提出了的新冠死亡率的序参量动力学模型，将新冠死亡率的复杂演变解耦为两个关键过程：早期阶段非幸存者的 ICU 持续时间和后期阶段的饱和死亡率。这一分析发现了新冠死亡率中个体尺度和国家尺度之间存在的跨尺度关联关系，从而

使我们能够定量地评估国家干预措施的效果，并且在疫情早期预测整个疫情的持续时间。这个模型框架也有其他潜在的应用，例如为国家对抗下一波疫情提供定量的建议，从而可以避免更多人的死亡。



## 第五章 应用于人体脉搏波系统

前文是将人体序参量理论应用于群体的演化，本章将继续把这一理论的应用由群体拓展到个体，开展对于人体脉搏波系统的分析。本章将在回顾总结以往研究的基础上，开展对于人体脉搏波的对称性分析并且提炼多层次的序参量，由此建立真正符合中医脉诊原理的序参量模型，为下一章刻画人体健康系统奠定基础。

### 5.1 脉搏研究的历史

在各种人体波动信号（如心电、脑电等）中，脉搏波信号由于十分方便获取，所以有着最为悠久的研究和应用历史。早在两千多年前，中国人就通过手指按压颈部、手腕、脚踝等部位的浅表动脉，感受其搏动情况来进行疾病诊断，并且绵延不断，至今仍然在临床一线服务，积累了丰富的实践经验，发展出一套系统的脉诊理论。

因此，从人体脉搏波信号入手，开展人体复杂系统的研究，具有丰富的理论基础优势。但是，脉搏波分析仍然面临着复杂性的挑战。本文认为，梳理脉搏波研究的历史，观其发展脉络，从中总结经验教训，对于我们深入理解脉搏波信号的复杂性，开展研究方法的创新，是十分有益且必要的。

#### 5.1.1 中医脉诊的研究

早在中华文明三大经典之一的《黄帝内经》（约公元前5世纪）中，就已经形成了较为完善的中医脉诊理论、方法和技术，并且明确地总结了人体脉搏波随着体质、情志、昼夜、季节、气候等变化的规律。《黄帝内经》阐述了四十余种脉象，并以“缓、急、大、小、滑、涩”作为统领诸多脉象的纲领<sup>[127]</sup>。

稍晚一些时间的《难经》（约公元前5世纪），提出了至今中医师主要使用的脉诊方法，即在手腕桡骨茎突处桡动脉上的“寸”“关”“尺”三个部位进行脉诊，每个部位又分别需要在“浮”“中”“沉”三个深度上开展测量。值得一提的是，《难经》十分重视脉诊的客观定量化，提出了以大豆（菽）重量为单位来量化记录取脉压力，在“浮中沉”不同深度上获取信息。

东汉的张仲景（公元150-219年）在经典著作《伤寒论》中，提出了“脉证并治”的诊断纲领，需要将“脉”与“证”相结合来辩证地诊断和治疗，即所谓“观其脉证，知犯何逆，随证治之”。《伤寒论》提到了约二十余种脉象，并且在开篇就提出了以“阴阳”为纲的脉诊纲领，“凡脉大、浮、数、动、滑，此名阳也；脉沉、涩、弱、弦、微，此名阴也”。

西晋的王叔和（公元 201-280 年）撰写了第一部较为系统的脉学专著《脉经》，梳理总结出了浮、芤、洪、滑等二十四种脉象，并且对每种脉象进行了描述。《脉经》还精确地描述了取脉的“寸”“关”“尺”的部位，阐述了不同取脉位置的脉搏波信号与反映的身体脏腑信息之间的对应关系，由此建立了统一的脉诊理论框架和操作规范，对后世影响很大。

此后的医家基本都是在《脉经》的基础上，结合自己的临床经验，提出了各自的脉诊主张。特别地，有两方面的变化值得注意：

第一，开始出现以“歌诀”的形式的脉学著作。例如，被李时珍编入《濒湖脉学》的《紫虚脉决》（公元 1189 年），就是以“四言”歌诀的形式总结了脉诊的理论要点。其言辞晓畅易懂，朗朗上口。这一现象说明，人们对于脉搏波与人体健康之间的关系已经有较为深入和系统的认识，所以才能够用直白的语言和方式，进行广泛流传。

第二，开始出现以“图画”的形式的脉学著作。例如，南宋施发撰写的《察病指南》（公元 1241 年）第一次收录了 33 幅脉象的脉图，用比较形象的方式来总结不同脉象的特点。虽然这时的脉图并不是脉搏波的客观记录图，但也反映了当时的医家已经开始尝试通过形象化的方式来表示不同的脉象。

通过以上的简要论述，可以看出在几千年的临床经验积累中，中医脉搏波的研究逐步形成了“提炼脉搏波要素，归纳典型脉象”的理论特点。所谓“要素”，是指最能够反映脉搏波运动特点的量。表 5.1 总结了不同时期的中医经典提炼的脉诊纲领，它们都是从不同的角度对脉搏运动特点的刻画。

表 5.1 不同时期的中医经典著作提炼的脉诊纲领

经典著作	年代	脉诊纲领
《灵枢》	先秦	缓急、滑涩、大小
《难经》	战国	浮沉、滑涩、长短
《伤寒论》	汉代	浮沉、滑涩、迟数
《濒湖脉学》	明代	浮沉、迟数、虚实

对于非专业读者，可能不太明白上表所罗列的“脉诊纲领”的具体含义，但仅从字面意思上来看，这些词语有一个共同的特点：它们都是由两个意思相对的字所组成的。事实上，它们描述的是某些关键量的“过甚”或“不及”的两种状态。例如，脉率过慢，则为“缓”；脉率过快，则为“急”。

进一步，基于脉搏波要素的不同组合，归纳出一些典型的脉搏波运动模式，称之为“脉象”。而这些典型的脉象，是进一步在临床开展辨证论治的基础，具有重要的临床意义。例如，南宋崔嘉彦的《紫虚脉决》提出以“浮沉、迟数”为纲，结合“脉力”的

“虚实”，统领了二十七种不同的脉象。明末李中梓的《诊家正眼》又在此基础上增加了“疾脉”，至此形成了现行中医院校统编教材《中医诊断学》中的二十八种脉象。

许多近代学者在继承传统脉诊方法的基础上，也提出了不同的脉诊纲领。例如，山东中医药大学齐向华教授从脉体、脉管、脉搏波、血流四个方面总结了 25 对脉象要素，构建了“系统辩证脉学”<sup>[128]</sup>。

综上所述，中医脉诊在几千年的实践发展中，形成了一套较为系统的理论。虽然在定量方面还不是很精细，但是通过在实践中总结出的刻画脉搏波运动的关键量，以及由此总结出的典型运动模式，就可以在临床中开展个性化的诊断。而这也正是中医脉诊的独特优势所在。因此，这一思想对于我们开展脉搏波分析的方法论创新，实现个体健康态的个性化辨析，有重要的指导意义。

### 5.1.2 现代脉搏波的研究

现代脉搏波的研究，是从人们对于血液的研究开始的。1628 年，Harvey 在 Michael Servetus 提出肺循环的基础上，提出了血液循环。1731 年，Stephen Hales 第一次在马身上测量了动脉血压，并基于蜡模第一次较为准确地计算了心脏输出量。1855 年，Karl von Vierordt 发明了脉搏描记器（Sphygmography）（见图 5.1），能够第一次无创地客观记录人体脉搏波<sup>[129]</sup>。

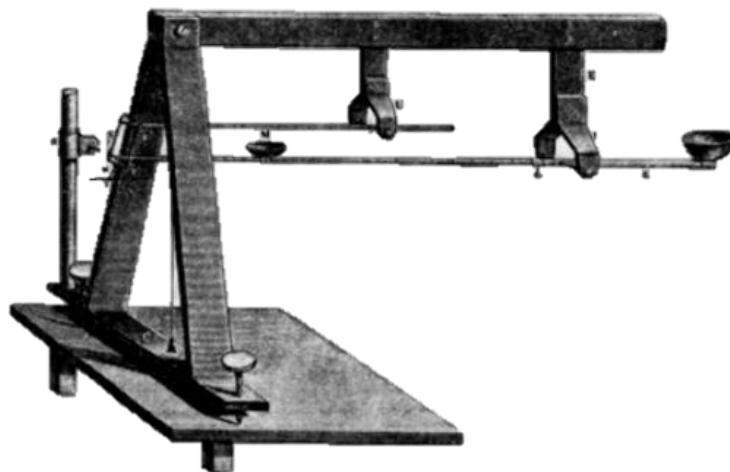


图 5.1 Karl von Vierordt 发明的脉搏描记器<sup>[129]</sup>

但是，随着 1896 年 Scipione Riva Rocci 发明肱动脉袖带水银血压计，以及 1905 年 Korotkoff 发现“柯氏音”，人们把注意力逐渐转移到了脉搏波搏动中的“收缩压”和“舒张压”上面。因为血压测量十分简便有效，在此后的一百多年里，“血压”成为了重要

的临床指标。并且伴随着电子血压计的发展，血压测量从临床走向了家庭。有学者认为，这同时也导致了“脉搏描记法和脉搏波分析从临床医疗中的消失”<sup>[130]</sup>。这一历史过程，值得我们今天开展脉搏波分析的研究者思考。正是因为人体信号的复杂性，所以必须要与临床之间形成有效的互相迭代，才能不断促进对于人体信号的深入理解。

另一方面，18世纪和19世纪，基于 Euler、Daniel Bernoulli<sup>①</sup>、Navier、Poiseuille、Stokes 等人的杰出工作，流体力学的发展推动了人们对于血液流动和脉搏波传播的认识<sup>[131]</sup>。1899年，Otto Frank 等人尝试将血管系统抽象为一个有输入和输出的弹性腔，建立了弹性腔模型来模拟脉搏波的传播。为了模拟更为真实的血液动力学系统，学者们在线性弹性腔模型的基础上不断拓展，例如形成了许多包含了非线性、多腔体的模型。但是由于血管系统的空间异质性、血液流动系统的复杂性，弹性腔理论在临床应用中遇到了较大的挑战。

然而，令人启发的是，在临幊上得到广泛应用的，却是脉搏波传播中的一个宏观变量——脉搏波速度（Pulse Wave Velocity, PWV），它已经成为了临幊检测血管僵硬度的“金标准”<sup>[132]</sup>。

根据两个测量部位之间的距离除以脉搏波到达的时间差，就可以很方便地计算出脉搏波速度（PWV）。由于血液是不可压缩流体，能量传播主要发生在动脉壁上，因此脉搏壁的弹性、厚度和直径等特性是影响 PWV 的主要因素<sup>[130]</sup>。因此，1808 年，Thomas Young 将脉搏波在动脉中的传播类比为声音在固体或不可压缩流体中传播，推导出了脉搏波传播的公式，也正是在这个过程中提出了著名的“弹性模量”的概念。以后，Moens、Bramwell、Hill、Bergel 等人又根据实验数据进行了模型修正。目前，脉搏波速度在预测连续血压、评价心血管危险因素、评估预后等方面有着广泛的应用<sup>[130]</sup>。

上述这段简要的历史，引发我们思考：为什么基于微观动力学的建模遇到了很大的困难，而作为宏观变量的脉搏波速度却在临幊评价中成为“金标准”呢？这是一个值得深思的问题。

从人体序参量理论的视角来看，脉搏波速度就是一个表征心血管系统有序性的序参量。虽然微观的血流系统十分复杂，涉及到细胞、血管、组织和器官等之间的复杂耦合作用，但是在对称性的约束下，最终一定会产生一个有序结构，而脉搏波速度就是评价这个有序结构的宏观量。

## 5.2 脉搏波研究的挑战

根据脉搏波研究所涉及到的整个系统，学者们在脉搏波的波形采集、波形预处理、

<sup>①</sup> Daniel Bernoulli 在传染病模型方面也做出了开创性的工作，详见第三章相关内容

特征提取和分类器方面，开展了丰富的研究<sup>[133]</sup>。

图 5.2 展示了脉搏波信号的一个典型周期<sup>[133]</sup>。一般认为，动脉的脉搏波包含三个组成部分：主波、潮波（亦称“重搏前波”）和重搏波。一个单周期的脉搏波由上升沿和下降沿组成，主波是由上升沿和部分下降沿组成的。潮波和重搏波，即第一和第二个反射波，二者之间小凹陷，被称为“重搏切迹”。潮波、重搏切迹和重搏波组成了脉搏波剩余的下降沿部分。在重搏切迹前的部分反映了心脏收缩的状态，切迹后的波反映了外围的状态。

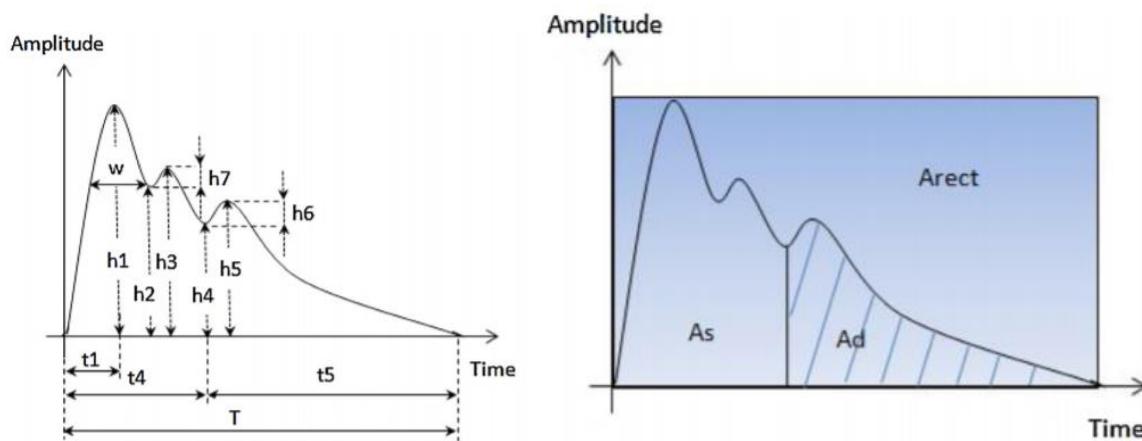


图 5.2 脉搏波信号的时域参数<sup>[133]</sup>

表 5.2 梳理总结了在目前脉搏波的研究中，经常用来表征脉搏波波形特征的一些主要参数<sup>[134]</sup>。

表 5.2 脉搏波信号的时域参数的描述

参数	描述	参数	描述
$h_1$	主峰振幅	$t_1$	脉搏波起始点与峰值之间的持续时间
$h_2$	主谷振幅	$t_4$	脉搏波起始点与切迹之间的持续时间
$h_3$	潮波振幅	$t_5$	重搏波起始与结束点之间的持续时间
$h_4$	切迹波振幅	$T$	一个脉搏周期的时间
$h_5$	重搏波振幅	$W$	主波在 $1/3 h_1$ 处的宽度
$h_6$	重搏波高度	$As$	收缩期面积
$h_7$	潮波高度	$Ad$	舒张期面积
-	-	$A_{rect}$	由 $h_1$ 和 $T$ 定义的矩形面积

### 5.2.1 脉搏波研究的现状

本文更侧重于关注脉搏波分析在应用方面的研究，所以根据研究目标的不同，将这些研究主要分为了两大类：一类是“中医脉诊客观化”的研究，一类是“生理病理分类”的研究。

#### 中医脉诊客观化的研究

上世纪 50 年代以来，人们在中医脉诊客观化方面开展了大量的基础研究。所谓“脉诊客观化”，就是要将传统的中医脉诊进行客观量化。例如，费兆馥等人总结了 70-80 年代 10 余年的脉诊研究成果<sup>[135]</sup>。李永光等人<sup>[136]</sup>从 2000~2008 年在国内对健康人群和病人进行了大样本的脉象普查，绘制了临床常见的 38 种常见脉象的脉搏波图，并且对每种脉象的脉搏波时域参数进行了统计总结。

在应用方面，这方面的研究主要是对不同中医脉象、证型的人群做分类。例如，Wang<sup>[137]</sup>发展了测定脉搏强度的方法，将脉搏强度分为正常、充满和微弱三类。Kim 等<sup>[138]</sup>提出了一种基于不同压下脉搏信号幅度信息的浮脉和沉脉的识别算法。Li 等<sup>[139]</sup>提出了一种区分滑脉和涩脉信号的方法。Xu 等<sup>[140]</sup>提出使用复杂性分析来区分三种心律不齐的脉象和四种心律不齐的脉象。Zhang 等人<sup>[141]</sup>指出，根据脉搏信号中主波、潮波、重搏波的脉搏峰值的存在、缺失和强度，可以定义五种形状相关的脉象，即缓脉、滑脉、紧脉、芤脉和涩脉。Guo 等<sup>[142]</sup>的目的是通过脉象诊断区分四种中医心律：心气虚、心阴虚、心阳虚、心脉阻痰。

上述研究，虽然在小样本下可以实现对于几种典型脉象的识别，但是由于真实人群的脉象往往是多种典型脉象的叠加，这大大制约了其在临床的可应用性。同时，由于不同医师的临床经验不同，对于脉象的划分标准不统一，导致对于同一个病人得出的脉象划分结果往往不一致。更重要的是，即使能够准确地判断出脉象，还是需要依靠医师来进一步诊断分析，因此在临幊上应用的价值有限。目前更多是作为辅助教学的工具在实践中使用。

#### 生理病理分类的研究

“生理病理分类”的研究，主要是对不同健康态、病理的人群做分类。例如，对于心血管疾病，参考文献<sup>[143-145]</sup>分析了腕部脉搏信号与高血压的相关性。Xu 等人<sup>[146]</sup>研究了评估冠状动脉硬化状态的脉搏波形。还有研究<sup>[147-151]</sup>旨在确定冠心病患者、非冠心病患者和健康个体之间的脉搏差异。另有其他研究是在区分两种健康态，如 Liao 等人<sup>[152]</sup>报道了不同孕期脉搏频谱的差异，Zhang 等人<sup>[141]</sup>分析了区分健康人与胆囊炎或肾炎患者的方法，Huang 等人<sup>[153]</sup>旨在鉴别消化不良和鼻炎，Gong 等人<sup>[154]</sup>应用了一种脉搏传感装置来识别肝硬化。Thakker<sup>[155]</sup>比较了胃肠道疾病患者和健康受试者的脉搏信息。

Zhang 等人<sup>[156]</sup>, 针对高血压、亚健康、肝硬化、冠心病和妊娠患者等不同健康态进行区分。

上述两类研究主要采取的是统计学习的方法, 通过对群体样本开展统计分析, 确定脉搏波的特征参数, 来对不同的状态——基本上只能对两种状态——进行区分。然而, 人体脉搏波受到个人体质、情绪、外界环境等多种因素的耦合作用影响, 呈现出复杂的时空异质性。当前基于统计学习方法的研究, 缺乏清晰的物理图像支撑, 无法定量刻画个体健康态, 从而面临着无法应用于临床的困境。

### 5.2.2 脉搏波研究的复杂性

通过上述对于脉搏波研究历史和进展的介绍可以看出, 脉搏波的研究中充满了复杂的挑战。笔者根据研究中的一些体会, 将研究中遇到的复杂性挑战总结为以下三个方面。

#### 脉搏波信号的复杂性

首先, 是来自于脉搏波信号的复杂性。这里又可以分为两方面, 一方面是脉搏波信号本身的复杂性, 比如信号的稳定性、敏感性、可重复性等, 这对于规范测量提出了挑战。另一方面则是由不同的测量条件带来的。

目前, 不同的研究采用的传感器、测量部位、测量方法不一样, 所得到的脉搏波信号有很大的差别。例如, 在手指、耳尖和手腕部采取的脉搏波的波形有很大的差异。所以, 当前研究自主采样获得的样本量都较小, 且不同的研究之间缺乏可比较的数据基础。目前虽然有像 MIMIC 等开源数据, 但相对于所研究问题本身的复杂度而言, 其样本量和覆盖面还是很不足的。

本文认为, 应积极推动国家层面出台相关测量标准, 开展不同地区、不同人群的数据采集, 逐步建立起人体脉搏波信息数据库。这对于深入开展人体脉搏波信号的研究, 具有十分重要的意义。

#### 数据标记的复杂性

其次, 是来自于数据标记的复杂性。目前的研究主要是二分类的问题, 例如通过脉搏波信号来区分患者某疾病的“患者”和“健康人”, 这里的“患者”虽然可以通过病理检查来确定, 但是同时还要尽量排除其他疾病——这并不是一件容易的事情。而“健康人”的定义往往又是模糊不清的。的确, 在很多研究文献中也并没有给出明确的纳入标准。而对于“中医脉诊客观化”的研究, 则需要人为来标记“脉象”“证型”等。但不同中医师的理论、经验、技法等不同, 对同一个病人得出的结论也往往不相同。

本文认为, 人体的生理状态是一个多维多层次的有序状态。因此, 对人体健康情况

的标注，需要建立这样一套多维多层次的标注体系。本文构建的《稳定健康指数量表》和《日变健康指数量表》，就是在这方面进行的初步尝试。

### 结果解释的复杂性

最后，是来自于结果解释的复杂性。这里涉及到不同层面的问题。第一，如前所述，由于缺乏公共数据的基础，不同研究结果之间的可比性大大降低。第二，目前主要采用的统计学习方法，对结果的主要的评价指标是分类的“精度”。但是在有限样本下，相对于人体复杂性而言，结果精度比较的可靠性大大降低。第三，目前研究结果对于临床应用的价值较低，比如通过已有的医疗仪器就可以精确地区分不同病症，这就使得研究和实践之间无法形成良好地互动和迭代，也就不能促使研究逐步深入进去。

本文认为，人体脉搏波信号，相比机器学习任务的传统对象，例如文字、语音、图像等信号，具有明显的个体性和强涨落等特征。因此，不能简单地将这些方法应用于人体脉搏波信号。在目前阶段，相比更为精准地分类，更加需要的是对人体脉搏波信号开展更为深入的理解。

综上所述，本文认为，当前脉搏波研究的挑战在于对人体脉搏波信号的深入理解。几千年的中医脉诊实践证明，脉搏波信号中蕴含着丰富的人体系统的多维多层次的信息。而中医脉诊的优势也正是在于能够有效地理解和应用这些信息，从而开展个性化的诊断。

因此，中医脉诊客观化研究不仅是要对不同“脉象”或“证型”进行区分。更重要的是，要对于人体脉搏波信号的理解做出重要贡献。而这，恰恰也是“中医脉诊客观化”研究与一般的脉搏波研究所不一样的重要区别。具体来说，需要从中医脉诊理论出发，构建真正符合中医脉诊理论特点的脉搏波分析模型，对个体的脉搏波信号开展从定性到定量的深入分析，总结人体脉搏波信号的普适规律。这不是一件容易的事情，需要方法论的重大创新。这也是本章所要尝试开展的研究内容。

### 5.3 脉搏系统的多层次对称性刻画

面对上述挑战，本文将人体序参量理论应用于脉搏波分析中，希望对人体脉搏波的信号开展更为深刻的理解。首先，就是要对人体脉搏波传播过程中的对称破缺过程开展分析。但是，脉搏波的传播涉及到人体多个组织、器官、系统等复杂的相互作用，是一个十分复杂的动力学过程。因此，本文将关注的重点放在与辨析人体健康态有关的过程上。

### 5.3.1 个体常态的对称性刻画

中医理论中提出了“平人”的概念和思想。所谓“平人”，可以理解为是一个“标准健康”的模型人。通过对比个体与“平人”的状态，就可以来判断、调理病情。例如，《黄帝内经·平人气象论》中提出，“平人者不病也，常以不病调病人”。根据“平人”的思想，又进一步提出了“平脉”的概念，即“标准健康”状态的人体脉搏波。平脉的特点是一种既不“过甚”，也没有“不及”的状态。

但是，这其实是一种比较理想的状态。实际上，在体质、情志、环境等多种复杂因素的长期作用下，个体脉搏波的运动状态无法始终保持在绝对的平衡态，而总是会发生某些偏离，并且最终形成个性化的脉搏波运动的“常态”。

例如，就以最容易辨析的脉搏波的搏动位置（脉位）和搏动力量（脉力）来说。对于身体质量指数（BMI）较高的个体，他们的脉搏波的搏动位置一般远离表皮而更接近桡骨，即脉位较“沉”；而 BMI 较小的个体的脉位则相对较“浮”。同样，在临床中发现，男性的脉搏波搏动的主波振幅较大，即脉力较“实”；而女性的脉力则相对较“虚”。这里所提到的脉“沉”“浮”“实”和“虚”等，并不是一种病理性的判断，而是反映了个体的脉搏波运动的一种常态特征。

从对称性分析的角度对上述讨论进行理解，“平人”的脉搏波的特点是不沉不浮、不实不虚的，所以具有一种对称性。但是在先天和后天、内在和外在等多种因素的长期作用下，个体脉搏波的运动发生了对称破缺，逐渐形成了一种“常态”，即新的有序性。

下面，我们通过一个简单的例子来解释。例如，《医学传心录》中在讲解不同的脉象时提到，“浮而无力即为虚，浮而有力便为洪。沉而无力即为弱，微沉有力是为实”。如图 5.3 所示，如果建立一个由“脉位”和“脉力”所组成的相空间，平脉则处于这个相空间的相对平衡的位置。此时，整个相空间具有一种对称性。而当发生对称破缺后，如果用最简单的“过甚”或“不及”来定量划分，原来的相空间就可以简单地分为四种典型的有序状态：浮实、沉实、沉虚和浮虚。这些不同的有序状态，分别反映了重要的人体健康态信息，在中医脉诊中具有非常重要的临床应用价值。而这里的“脉位”“脉力”就是重要的序参量，因为它们表征了对称破缺后涌现出的新的有序性。

需要说明的是，上面所说的“平脉”，并不是一个恒定不变的绝对状态，而是会随着外界环境变化而相应变化。比如，《黄帝内经》中讲，春季“微弦曰平”，夏季“微钩曰平”，长夏“微弱曰平”，秋季“微毛曰平”，冬季“微石曰平”。这就说明，人体系统本身就随着外界四时的变化而变化，因此人体脉搏波的平衡状态也会随之而变。这种“天人相应”的系统观贯穿于中医脉诊理论和实践的始终，对于本文开展脉搏波的对称性分析和序参量建模，具有重要的启发意义。

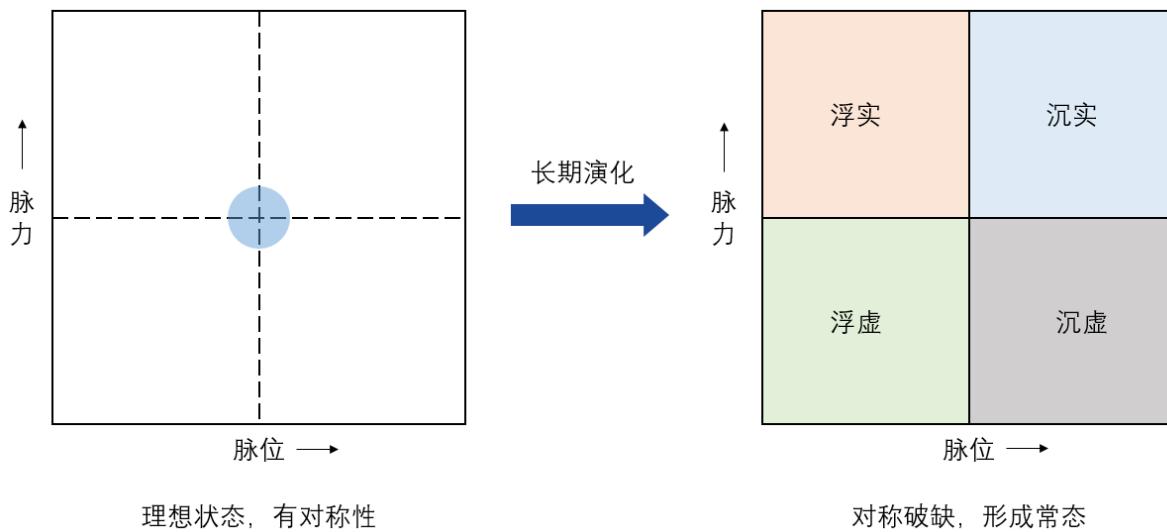


图 5.3 个体脉搏波的长期演化发生对称破缺形成常态

### 5.3.2 个体偏离态的对称性刻画

由于人体是一个高度开放的复杂巨系统，人体系统随时在与外界环境之间进行着物质、能量和信息的交换。因此，这就使得人体脉搏波信号在传播中再次发生对称破缺。从而，个体脉搏波的实际运动状态，并不是一直处于“常态”而不改变，而是时刻围绕着“常态”在波动，而形成“偏离态”。

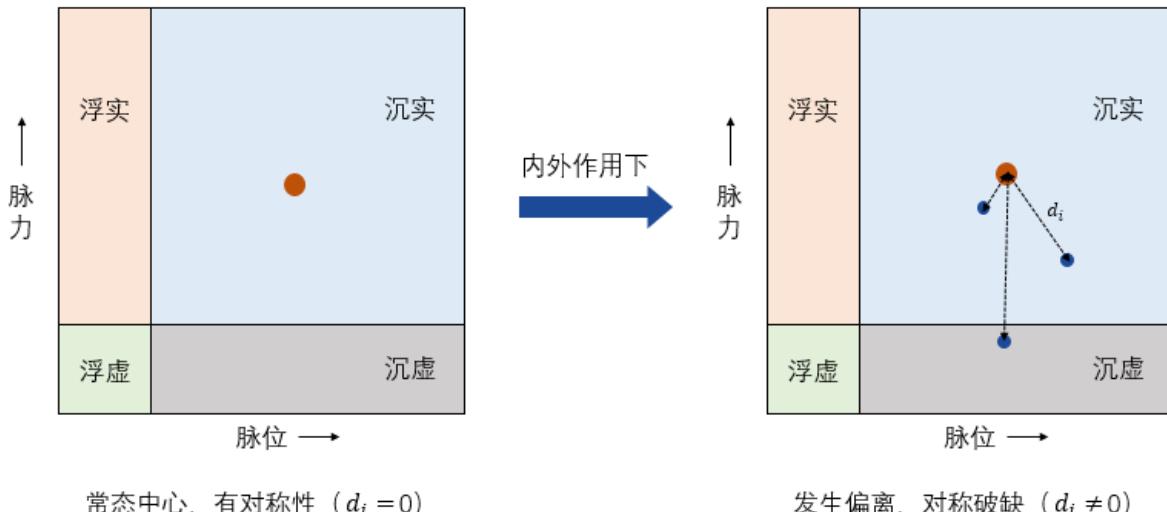


图 5.4 个体脉搏波在内外作用下发生对称破缺形成偏离态

如图 5.4 所示，个体脉搏波处于某一个特定常态时，这个常态的相空间具有对称性。但是在内外作用下，个体脉搏波会围绕这“常态”发生偏离，对称破缺，而处于“偏离态”。

本文定义偏离态与常态之间的距离为“偏离度”( $d_i$ )。当 $d_i = 0$ 时，系统没有发生对称破缺；而当 $d_i \neq 0$ 时，系统发生了对称破缺。因此，“偏离度”是刻画这个过程中涌现出来的有序性的序参量。

### 5.3.3 知常达变的序参量解释

中医脉诊通过确定个体的“常态”，并且根据当下状态与“常态”之间的偏离程度，来进行个性化的诊断。这被称之为“知常达变”的原理和方法。例如，在《黄帝内经·平人气象论》中指出，偏离“常态”的不同程度和不同情况，与健康状态之间有着密切的关系。以春季为例，“春胃微弦曰平，弦多胃少曰肝病，但弦无胃曰死；胃而有毛曰秋病，毛甚曰今病”。这句话的意思是说，在春天比较平衡的脉是弦中带有胃气。如果“弦”过多，则会有肝病；如果“弦”过盛，病情就很危急。而如果同时出现“毛脉”，则会秋季发病；如果“毛脉”过盛，则当下就会发病。

通过上文的论述，我们就可以从序参量的角度对“知常达变”进行更加定量的理解。如上所述，“偏离度”( $d_i$ )定量地刻画了个体当下状态与“常态”之间的距离。当偏离度 $d_i$ 处于一定阈值范围内时，个体状态属于正常的波动范围。而当偏离度 $d_i$ 超过一定阈值后，则处于异常的状态。例如，前文介绍的BMI较大的个体的脉位比较沉，不能因此就被判断为是异常状态。而是应该综合考虑其平时的脉位水平后，再做判断。如果平时脉位就很沉，说明“脉位沉”恰恰是该个体的“常态”。而如果平时脉位很沉，当下脉位忽然变得很浮，这说明当下处于异常状态。

因此，偏离度 $d_i$ 就是表征人体健康状态的序参量。通过确定个体的“常态”，并且度量当下状态与“常态”之间的偏离度 $d_i$ ，就可以开展个性化的个体健康状态的定量辨析。

综上所述，我们通过运用人体序参量理论，对脉搏波系统开展对称性分析，发现了脉搏波系统中存在着多层次的对称破缺过程。由此，我们建立了基于“常态”和“偏离度”的个性化辨析健康状态的新方法。下面，我们将据此开展人体脉搏波序参量模型的建立。

## 5.4 提取多层次的脉搏波特征参数

为了建立人体脉搏波的序参量模型，我们首先要开展多层次的脉搏波特征参数的提取工作。

### 5.4.1 实验测量与数据采集

如上文分析，人体脉搏波存在自身节律，并且很容易受到外界环境的影响。因此，

我们在充分地考虑这些特点的基础上，设计了两类脉搏波测量实验。

第一类是日常测量实验。实验员每天在固定时间、固定地点开展脉搏波的测量。测量前和测量时，需要保持在静息状态。测量实验结束后，需要填写《日变健康指数问卷》（详见 6.1 节），对当日健康情况进行标记。由此，就形成个体的脉搏波的长期连续测量数据集。这些数据十分宝贵，对于揭示人体脉搏波的波动状态和节律特征，具有十分重要的意义。同时这也是本研究的一大特色。

第二类是群体测量实验。经过培训的实验员，在社区随机采集不同年龄段人群的脉搏波数据。在实验结束后，实验员辅助被测者填写《稳定健康指数问卷》（详见 6.1 节），来对受测者的健康情况进行标记。由此，就形成了群体的脉搏波测量数据集。

具体来说，本研究使用了一款由步进电机驱动的单点加压式、压电传感式脉搏波采集器（见图 5.5），其采样频率为 50Hz，采样时间为 10s。每次完整的测量，需要分别在左右手的手腕挠动脉处的寸、关、尺共六个部位（见图 5.5），采集压力脉搏波的压力信号。



图 5.5 脉搏波的采集仪器和测量部位。左图为脉搏波采集仪器<sup>[157]</sup>，右图为六个测量部位。

在每个部位测量时，脉搏波采集器将分别在反馈压力为 75g、100g、125g、150g、175g、200g、225g 和 250g 左右时，采集 10 秒钟的脉搏波信号。这样在每个测量部位，就会有在 8 段压力下的脉搏波数据，如图 5.6 所示。

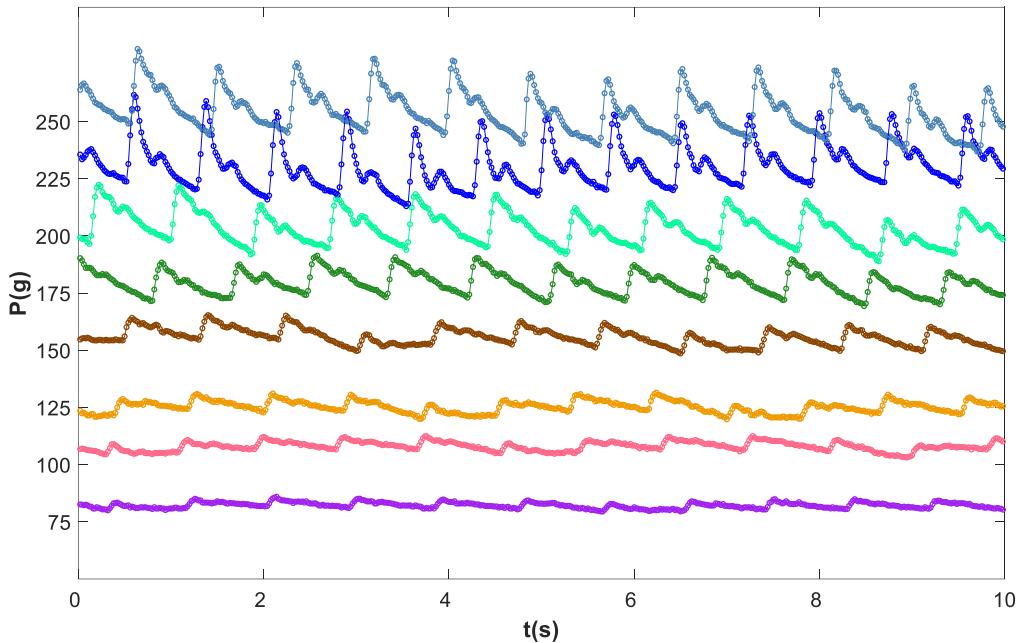


图 5.6 在 8 段压力下的脉搏波原始信号。

#### 5.4.2 数据预处理

##### 周期分割

首先，针对每段加压压力下的脉搏波信号（500 个采样点），进行窗口滑动寻找局部最小值的方法，确定出脉搏波的每个周期的起始点（图 5.7 中的红色圆圈）。而每个周期的起始点，又是上一个周期的结束点，从而实现对每个周期的分割。

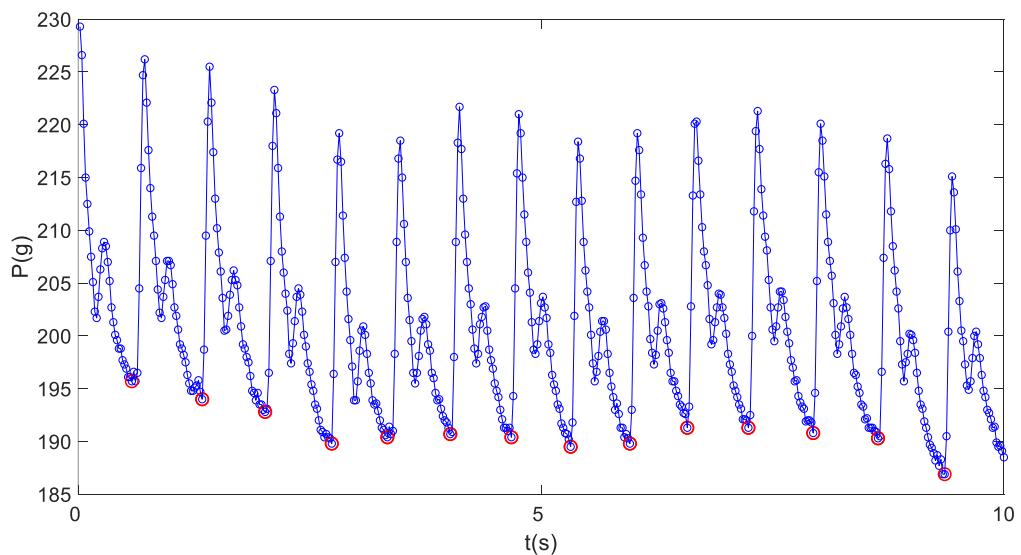


图 5.7 对每段脉搏波进行周期分割。

### 去除基线漂移

对于上一步分割出来的每个完整的周期，需要去除基线的漂移。这里使用了常用的方法：对原始数据减去起始点和结束点的连线上的值，从而使得起始点和结束点的值都为 0，实现了基线漂移的去除，如图 5.8 所示。

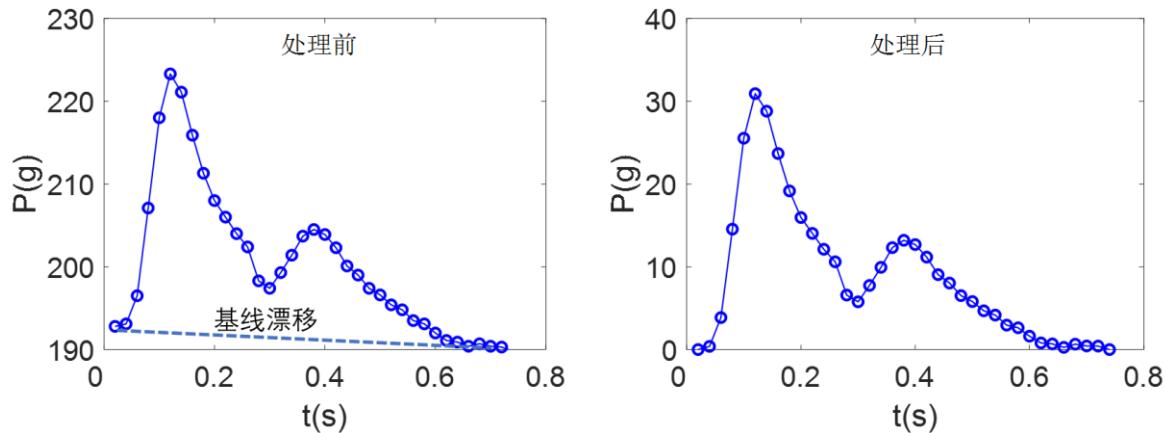


图 5.8 基线漂移去除前后对比图。

### 确定最佳脉搏波数据

对于每个周期的脉搏波数据求最大值，就确定出了该周期的主波振幅  $f_i$ （见图 5.9）。对某压力下的多个周期的主波幅值  $f_i$  求平均，则得到在该压力下的平均主波振幅  $F_j$ ，即：

$$F_j = \frac{\sum_i^n f_i}{n} \quad (5.1)$$

这里， $j = 1, 2, \dots, 8$ ，为 8 段压力的标记； $i = 1, 2, \dots$ ，为每段压力下的各周期的标记。 $n$  为在 10 秒采样数据中的完整搏动周期的个数。

进一步，我们选择 8 段压力下的主波幅值  $F_j$  为最大时的压力为最佳取脉压力，而该压力下的脉搏波数据，就是最佳脉搏波数据。

对每次测量时的六个部位，都进行上述步骤的预处理，则得到六个测量部位的最佳脉搏波数据。

#### 5.4.3 主导特征参数

中医脉诊理论中将“脉位”“脉力”和“脉率”，作为辨析脉搏波状态的最基本的三个量。例如《医学传心录》中讲，“诊家之要四般脉，浮沉迟数为之则”。为此，本文提出了相应的“基线压力”“主波振幅”和“搏动周期”三个可测量，如图 5.9 所示。

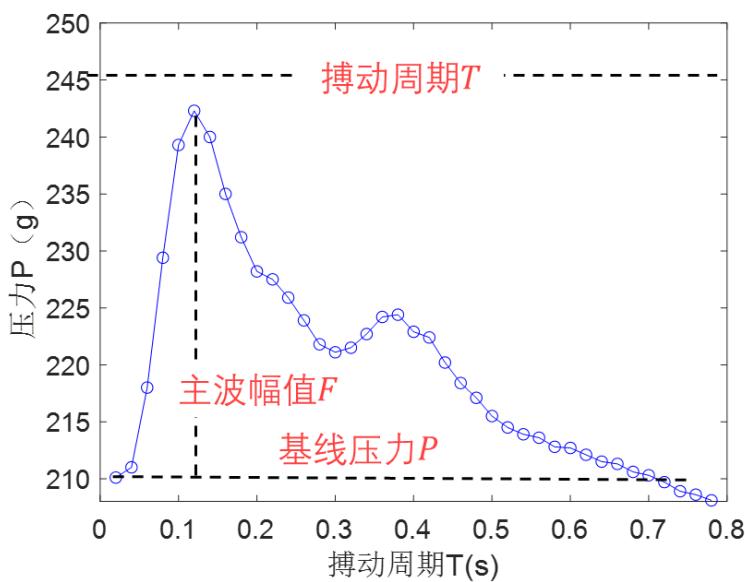


图 5.9 脉搏波主导特征参数的提取。

### 脉位与基线压力 $P$

在对脉管的径向进行加压的过程中，脉管所反弹的压力大小随着加压的深度而改变。如果脉管主要在表皮附近搏动，则为“浮”的状态；如果脉管主要搏动的位置更接近桡骨，则为“沉”的状态。因此，脉位可以由基线压力确定，刻画了脉搏波传播的空间维度的主要特征。根据中医脉诊理论，脉位的不同状态，与人体气机的表里程度等密切相关<sup>[135]</sup>。因此，“脉位”是重要的脉搏波特征参数之一。

### 脉力与主波振幅 $F$

根据脉管所反弹的压力的大小，可以分为“虚”与“实”两种典型状态。如果压力过小，则为“虚”的状态；如果压力过大，则为“实”的状态。因此，脉力可以由主波振幅确定，刻画了脉搏波的振幅的主要特征。根据中医脉诊理论，脉力的不同状态，与人体气机的虚实程度等密切相关<sup>[135]</sup>。因此，“脉力”是重要的脉搏波特征参数之一。

### 脉率与搏动周期 $T$

如前所述，脉率过慢则为“迟”，而脉率过快则为“数”。所以，脉率可以通过脉搏波的搏动周期来确定，刻画了脉搏波传播的时间维度的主要特征。根据中医脉诊理论，脉力的不同状态，与人体气机的寒热程度等密切相关<sup>[135]</sup>。因此，“脉率”是重要的脉搏波特征参数之一。

上述“基线压力”“主波振幅”和“搏动周期”，对脉搏波的传播中的最主要特征进行了刻画（见图 5.9），因此本文称之为脉搏波的“主导特征参数”。

## 六部平均值

进一步，我们对六个测量部位的脉搏波的主导特征参数，分别求平均，得到六部的平均值：

$$\bar{P} = \frac{\sum_1^6 P_i}{6} \quad (5.2)$$

$$\bar{F} = \frac{\sum_1^6 F_i}{6} \quad (5.3)$$

$$\bar{T} = \frac{\sum_1^6 T_i}{6} \quad (5.4)$$

这里， $P_i$ 、 $F_i$ 和 $T_i$ 分别是基线压力、主波振幅和搏动周期。 $i = 1,2,3,4,5,6$ 分别代表着左寸、左关、左尺、右寸、右关、右尺这六个测量部位。

## 长期测量结果

这里报告一位长期测量人员（女，29岁），在2021年1月-2022年1月的测量数据（N=296）。这也是首次公开报道的个体长达12个月的连续测量结果。

为了更好地进行可视化展示，我们将其画在了3个二维相空间上，如图5.10所示。

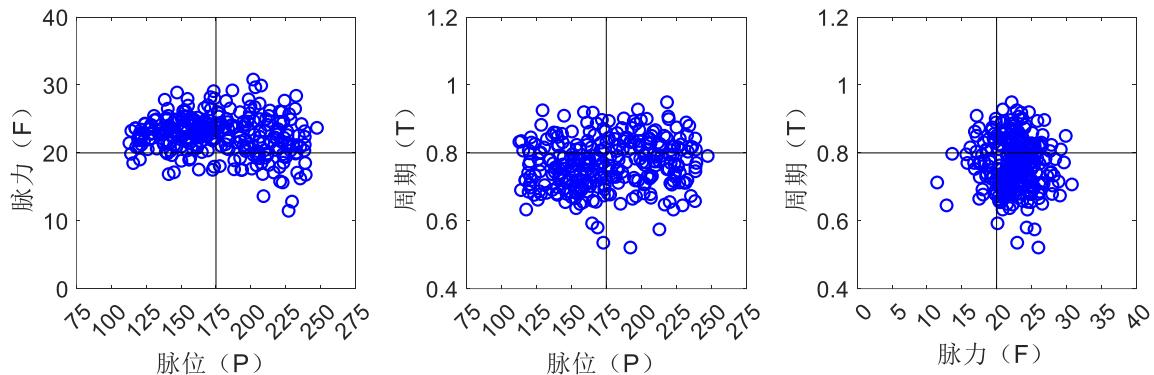


图5.10 连续12个月测量的脉搏波主导特征参数的六部平均值在相空间中的分布(N=296)。脉位为基线压力(单位：克)，脉力为主波振幅(单位：克)，周期为搏动时间(单位：秒)。

由上图可以看出，在长达12个月的个体在长期测量结果中，其主导特征参数的六部平均值在相空间中的分布是比较聚集在特定区域的。即该个体的脉力整体比较强，脉率整体偏慢，但脉位的变化范围较广。这验证了前文所述的个体脉搏波存在“常态”的认识。

同时，我们也可以看到，个体的脉搏波总是围绕着“常态”发生偏离的。特别是脉位的弥散度相对脉力和脉率而言要更加大。经过研究发现，这主要是跟其存在的亚日节律有关。

## 亚日节律特征

图 5.11-5.13 分别展示了三个主导特征参数的六部平均值在 2021 年 1 月-2022 年 1 月间 12 个月的演化行为。

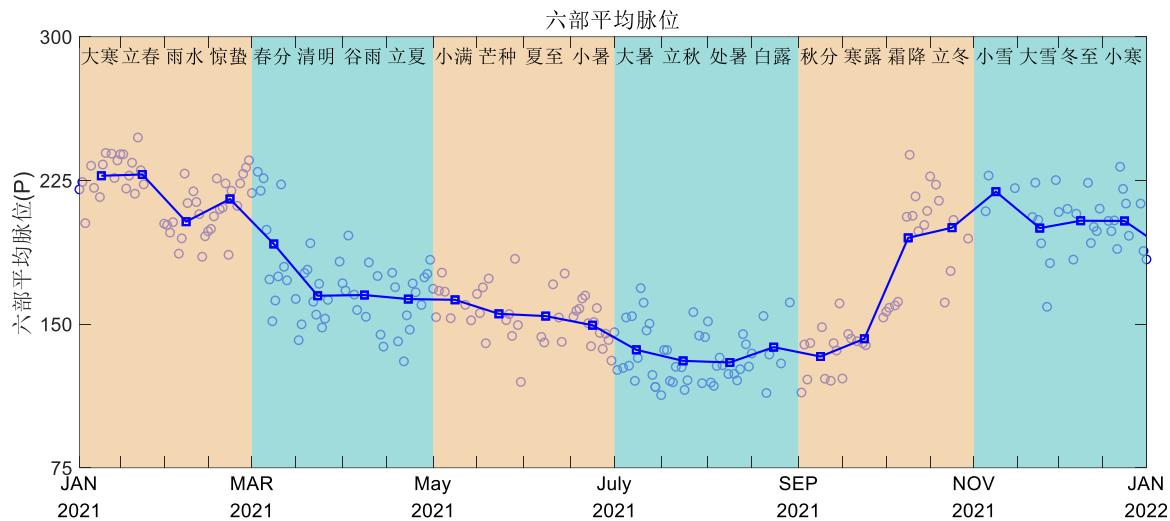


图 5.11 脉位的六部平均值（基线压力，单位：克）随时间的演化 其中圆点为每日测量值，方形点为一个节气脉位的平均值，不同的颜色为按照中医运气学说划分的时间节点，从左到右分别为第一至六气，每一气包含了四个节气。

首先，可以看到六部平均脉位在一年中有着较为明显的连续变化。秋冬季的脉位值较大，而春夏季的脉位值较小。比较有意思的是，脉位存在几个较大变化，其时间节点似乎与节气有一定的关系。例如，在“春分”前后，脉位开始明显变浮（小）。在“秋分”前后，脉位开始明显变沉（大）。而在其他时间，脉位都比较稳定。这说明，脉位很可能是一个与外界气候等有密切变化的物理量。

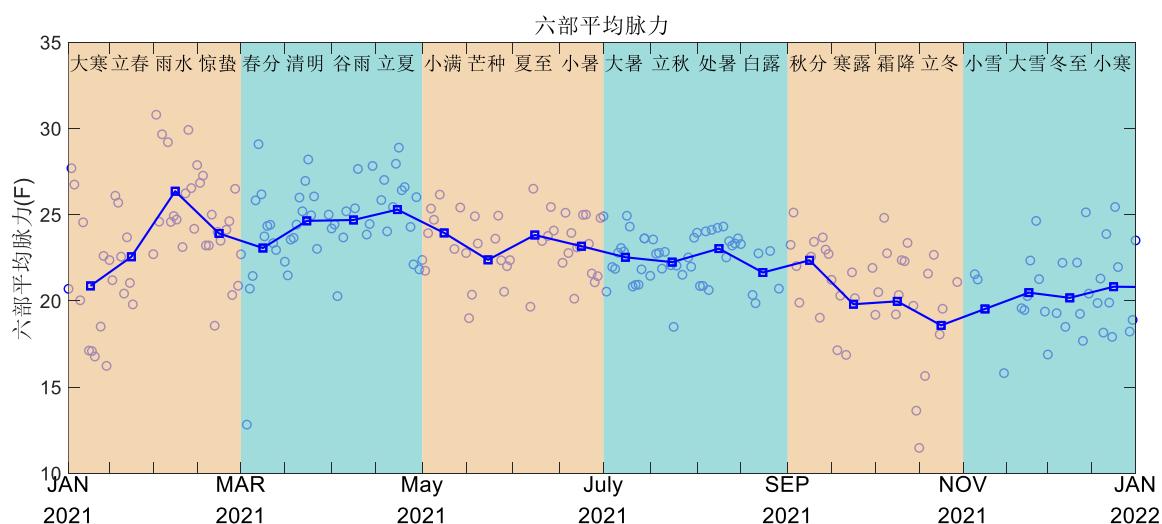


图 5.12 脉力的六部平均值（主波幅值，单位：克）随时间的演化 其中圆点为每日测量值，方形点为一个节气脉位的平均值，不同的颜色为按照中医运医学说划分的时间节点，从左到右分别为第一至六气，每一气包含了四个节气。

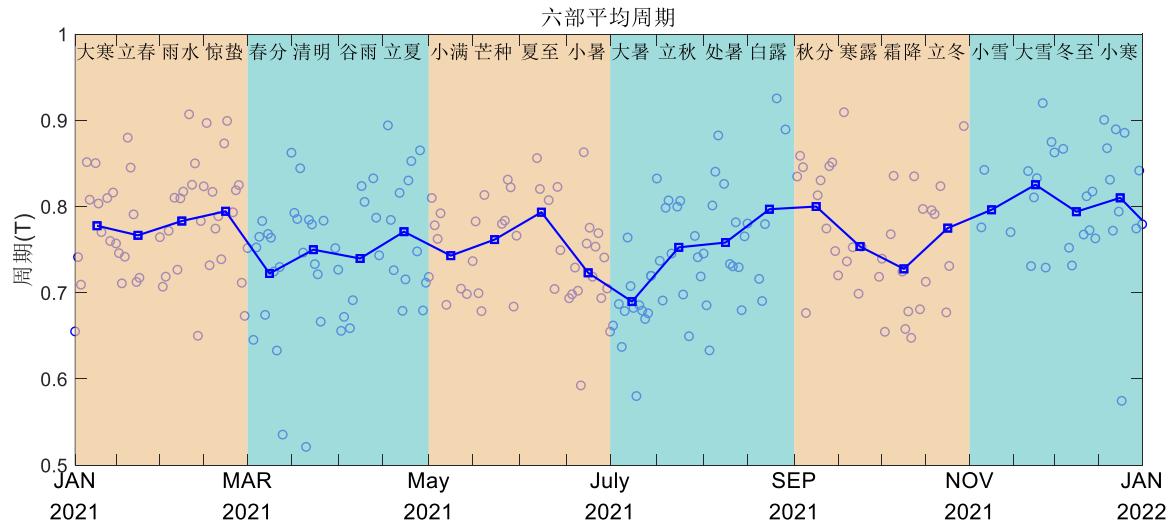


图 5.13 周期的六部平均值（搏动周期，单位：秒）随时间的演化 其中圆点为每日测量值，方形点为一个节气脉位的平均值，不同的颜色为按照中医运医学说划分的时间节点，从左到右分别为第一至六气，每一气包含了四个节气。

相对而言，六部平均脉力的节律特征不是很明显，常年相对比较稳定。而六部平均周期则涨落很大，这与人的心率很容易受环境影响而发生较大脉动的事实是一致的。通过上述简单的讨论，可以看出人体脉搏波的特征参数可能反映了不同时间尺度的人体系统的变化情况，值得在未来进一步研究。

#### 5.4.4 脉形特征参数

前述脉搏波的主导特征参数刻画了脉搏波传播中的主要特点，而更多的信息则蕴含在脉搏波的波形特征中。由此，我们提出了高斯函数模型来确定脉搏波的波形特征参数。

前人已经发现，多种人体生理波动信号，如神经电位、心电、脉搏波都具有高斯函数的形式，可以通过多个高斯函数叠加进行精确地拟合<sup>[158-162]</sup>。但是一直停留在拟合曲线的层面。

从对称性分析的角度，我们认为，不同来源的信号都具有高斯函数的形式，这说明这些信号传播的过程中，发生了类似的对称破缺的过程。因此，我们提出，用四个高斯波包的解来表征脉搏波的传播过程中的有序结构，如下式所示：

$$V_l = \sum_{i=1,2,3,4} a_i \exp \left\{ -\frac{(t - t_i)^2}{W_i^2} \right\} \quad (5.5)$$

其中,  $a_i$  为高斯波包的峰值高度,  $W_i$  是半高全宽,  $t_i$  是峰值的位置。

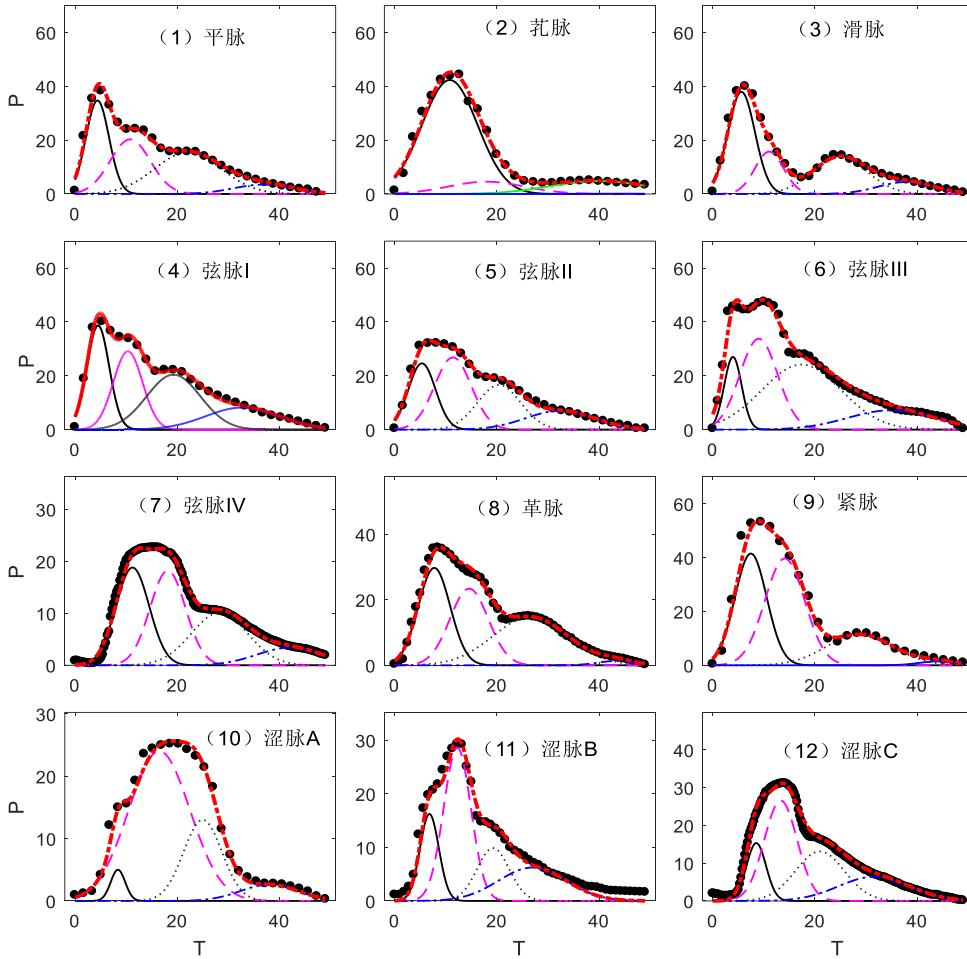


图 5.14 用四个高斯函数刻画多种典型脉象。

我们从《中国脉诊研究》一书中选取了 12 种典型脉象的脉搏波数据<sup>[135]</sup>, 进行了高斯函数的描述, 如图 5.14 所示。可以看出, 4 个高斯波包的模型, 能够精确地刻画多种复杂脉形的脉搏波。这些典型脉象在临床中具有重要价值, 反映了不同的人体健康状态。因此, 抓住了四个高斯波包的结构, 就抓住了脉搏波的简单性。四个高斯波包对应了人体脉搏波在传播过程中发生多次对称破缺的过程, 这对于理解人体生理和病理过程, 有重要的意义。

另外，相较于直接读图法，高斯函数对脉搏波的描述更具有稳定性，在信号噪音强或自身特征点不明显时，四个高斯波包依然能够抓住脉象的整体特征，保证参数的稳定性。

综上所述，本节提出了一套提取多层次的脉搏波特征参数的方法。首先，脉位、脉力和脉率是刻画脉搏波传播的主要特征的主导特征参数。通过报道并分析个体长达 12 个月的连续测量数据，我们发现主导特征参数具有明显的亚日节律特征。其次，针对更为丰富的脉搏波波形特征，我们提出了四个高斯波包解模型，成功对多种复杂的脉形进行了精确的刻画。由此，我们构建了一套多层次提取脉搏波特征参数的方案，为开展多层次健康态辨析的建模工作奠定了基础。

## 5.5 脉搏波分析的序参量模型

本节将在上述讨论的基础上，建立一个辨析人体健康状态的脉搏波序参量模型框架。这个模型框架第一次实现了将中医脉诊“知常达变”的原理和方法的模型化和定量化表达。下面来详细介绍。

### 5.5.1 序参量相空间

首先，我们提出了“序参量相空间”的概念。这一概念来源于中医脉诊理论中将不同脉搏波“要素”的组合并凝练成典型的脉象，并且在几千年的临床实践中，得到了验证（见 5.1 节）。

所谓“序参量相空间”，就是指在筛选出待研究问题的关键序参量  $O_n$  的基础上，由这些关键序参量所组成的相空间，就是序参量相空间  $\chi$ :

$$\chi = \{O_1, O_2, O_3, \dots, O_n\} \quad (5.6)$$

其中，下标表示不同的序参量。

根据人体序参量理论，不同的序参量是对人体系统涌现出来的新结构的有序性的刻画。那么，序参量相空间，则形成了对人体系统的有序性的多维度的综合刻画。例如，图 5.10 就展示了由三个主导序参量建立起的序参量相空间。

需要注意的是，这里的序参量相空间并不是唯一确定的，而是根据待研究问题所确定出的不同序参量，就构成了不同的序参量相空间。本文在第 6 章针对不同的研究问题，建立了不同的序参量相空间，均展示了良好的效果。因此，这一个概念具有极大的包容性和应用前景。

### 5.5.2 确定不同的亚类

根据上述讨论，序参量相空间中的不同位置，就代表了不同的有序状态。例如，对

于在图 5.3 中建立的“脉位-脉力”相空间，根据它的中心位置和围绕其周围的四个偏离位置，就可以将相空间分为“平衡态”“浮实态”“沉实态”“沉虚态”“浮虚态”共计五个典型的有序状态。而在同一类的数据，具有相同的有序特征。换句话说，具有相同有序特征的数据，在序参量相空间中应该聚成一类，我们称之为“亚类”。

当序参量相空间的维度很高时，由于数据量有限，我们无法先验地就确定出所有样本数据占据了该序参量相空间的哪些亚类，以及各亚类的范围。因此，我们选择采取了以数据驱动的方式来进行亚类的划分。

具体来说，我们使用了经典的 K-mean++的聚类方法，通过数据本身在序参量相空间中的聚集情况，来对样本聚类而实现亚类的划分。对于亚类数目  $K$  值的确定，可以通过数据本身的聚类评价效果以及模型的刻画效果来综合评价确定。由于 K-mean++是十分经典的聚类算法，这里不再赘述，感兴趣的读者可以进一步阅读参考文献<sup>[163]</sup>。

### 5.5.3 确定理想态和偏离度

如上所述，不同亚类代表了人体系统的不同的有序状态，因此不同亚类之间的健康状态没有直接的可比性。例如，脉力较强的男性，并不一定就比脉力较弱的女性要更加健康。但是，在同一个亚类内部，则可以开展有序度的比较。

假设每个亚类中都存在一个理想的健康状态，它的有序度最高，在序参量相空间中的位置可以用  $I$  来表示。在确定了理想态的位置后，我们就可以度量该亚类中的所有样本  $x_i$  与理想健康态  $I$  之间的距离，即偏离度：

$$d_i = \|x_i - I\| \quad (5.7)$$

其中，下标  $i$  代表了该亚类中的不同样本。

正如上一节所分析的那样，偏离度  $d_i$  是刻画人体健康状态的重要序参量。当偏离度  $d_i$  越大时，说明样本距离理想健康态越远，则有序度较低。反之，则有序度较高。

### 5.5.4 刻画人体健康数据

由此，就可以根据偏离度来刻画人体健康数据，而建立起二者之间关系的函数  $f(x)$ ，称之为激活函数：

$$H_i = f(d_i) \quad (5.8)$$

当激活函数为 logistic 函数时，就可以进行分类任务的刻画。而这里根据上述分析，反映人体健康水平的量，如上文提到的健康指数  $H$ ，与偏离度  $d_i$  之间是一个成反比的关系。假设二者之间是最简单的线性关系，则有：

$$H_i = k * d_i + H_{max} \quad (5.9)$$

其中,  $k$ 为反映脉搏波的偏离度变化对于健康水平影响的敏感因子。当 $d_i = 0$ 时,  $H_i = H_{max}$ 。所以,  $H_{max}$ 是理想态 $I$ 的健康指数。

我们使用最小二乘法, 使得模型预测的健康指数 $f(d_i)$ 与实际测量的健康指数 $H_i$ 之间的均方误差最小, 即可确定出理想健康态在序参量相空间中的位置 $I$ , 以及上述模型参数 $k$ 和 $H_{max}$ :

$$(I, k, H_{max}) = \arg \min \sum_{i=1}^n (f(d_i) - H_i)^2 \quad (5.10)$$

由此, 我们就完成了脉搏波序参量模型 (Order Parameter Model of Pulse Wave) 的构建。这里的 $H$ 是一个反映整体健康水平的变量。只要满足与偏离度之间的负线性关系,  $H$ 可以是来自于某个单一的指标, 也可以是多方面综合的结果。

### 5.5.5 模型的特点

首先, 总结一下建模过程中用到的基本假设, 主要为以下两点:

- (1) 序参量相空间中的每个亚类, 都有一个理想健康态 $I$ , 它的有序度最高, 拥有着该亚类中的最理想的健康水平。
- (2) 在每个亚类中, 样本 $x_i$ 与理想健康态 $I$ 之间的偏离度 $d_i$ , 与该样本的健康水平 $H_i$ 成反比。偏离度 $d_i$ 越大, 则距离理想态越远, 则健康水平越低。

其次, 脉搏波序参量模型的建模一般步骤, 如图 5.15 所示:

- (1) 根据待研究问题, 提炼关键序参量, 构建序参量相空间 $\chi$ ;
- (2) 根据数据在序参量相空间中的分布, 划分出不同的亚类;
- (3) 在每个亚类中, 度量样本 $x_i$ 与该亚类的理想态 $I$ 之间的距离 $d_i$ ;
- (4) 建立起偏离度 $d_i$ 与人体健康数据 $H_i$ 之间的函数关系;
- (5) 上述步骤多次迭代, 确定出各亚类的理想态 $I$ , 以及模型参数值。

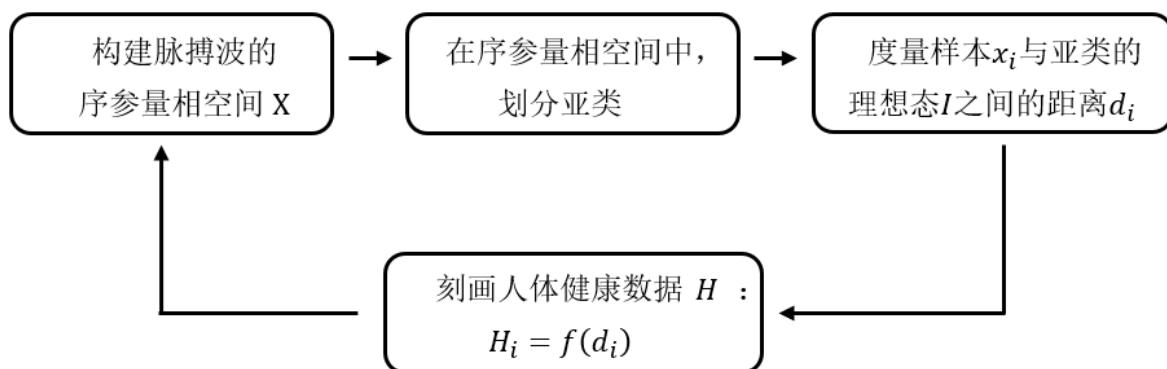


图 5.15 脉搏波序参量模型的建模一般步骤

最后，简要地总结一下脉搏波序参量模型的特点：

(1) 由于不同的待解决问题所提炼出的序参量是不一样的，所以这一模型框架具有十分大的包容性，适用于一大类的问题。

(2) 这一模型框架是“从定性到定量”的复杂系统研究方法的具体应用，并且进一步发展。即提出了“序参量相空间”的概念，并且基于“序参量相空间”来进行分类的方法。

(3) 这一模型框架将中医“知常达变”的原理和方法，进行了数理化的定量表达，突破了传统的基于统计学习的研究方法，建立了符合中医脉诊原理的原创新的脉搏波研究方法。

上述这些特点，在第6章的模型实际应用中得到了充分的展示。

## 5.6 本章小结

中医脉诊客观化经历了半个多世纪的发展，做出了许多重要的基础工作，但仍面临着无法应用于临床的挑战。本文认为，这需要突破当前主流的大数统计平均的方法局限，建立起真正符合中医脉诊理论的数理模型。本章将人体序参量理论应用于脉搏波的分析中，开启了这一尝试和探索。

本文基于对称性原理猜测，人体脉搏波系统在长期演化中形成一个个性化的“常态”，同时在内外作用下不断发生新的对称破缺，形成了对常态的偏离，由此凝练出“偏离度”这个关键序参量来定量刻画个体健康态的变化。这一认识与中医在几千年实践经验中总结出的“知常达变”的脉诊原理是十分符合的。

但是，如何将上述过程进行数理化的定量表述呢？本文提出了“人体序参量相空间”的概念。“人体序参量相空间”的不同位置，代表了不同的人体有序状态。因此，通过在序参量相空间中进行聚类，就可以识别出不同的人体有序态类型（亚类），这就构成了对亚健康人群划分的依据。在此基础上，通过确定出各亚类的理想健康态的位置，就可以计算出该亚类样本与理想健康态之间的偏离度。由此，我们就构建了脉搏波的序参量模型。

这一模型框架针对不同的待研究问题，可以建立不同的“序参量相空间”，显示了极大的包容性。为此，我们提出了一套包含脉搏波主导序参量和脉形序参量在内的参数提取方案。这为开展刻画不同健康状态的研究，奠定了重要的基础。

下一章我们将脉搏波的序参量模型，应用于多种实际健康数据的定量刻画，实现了对于当前大数统计平均研究的突破。



## 第六章 应用于人体健康指数的标定

本章将上一章构建的脉搏波序参量模型，应用于标定人体健康指数。分别在刻画稳定健康指数和日变健康指数方面，均得了数据的精确支持。本文还把这套模型框架，应用于刻画人体手指的气体放电可视化（Gas Discharge Visualization，GDV）信号，同样取得了成功。由此，验证了模型的有效性和普适性，并且显示了其对于人体波动信号的深入理解。

### 6.1 健康指数问卷

根据 5.4 节介绍，在完成脉搏波的测量后，需要对受测者的健康情况进行标记。但是，传统的以生理检测指标为主的标记方法存在着对人体健康状态进行过于碎片化和局部化标记的局限性（详见 5.2 节）。因此，本文选择通过建立多维多层次的问卷的形式，来对受测者的健康情况进行多维多层次的标记。

然而，目前已有的对人体健康开展测评的问卷，主要是针对人体健康的某一方面或某几个维度（如高血压量表）或者特定人群（如老年人健康综合评估量表）开展评价，缺乏对于人体系统的整体的、主客观统一的健康评价问卷或量表。为此，本研究团队以“人体复杂系统论”为指导，专门设计开发了一套《健康指数问卷》，用于对受测者的健康状态的标记。

由于这部分内容不是本文的重点研究内容，本文在这里进行一个相对简要的介绍，详细内容可以参考北京大学医学部刘奕君博士的相关论文内容<sup>[164]</sup>。

如前所述，人体是一个“一元二面多维多层次”的开放复杂巨系统，人体系统的健康状态围绕着理想的健康状态在不断波动。如果偏离理想态太远，就会体现出结构或者功能方面的症状。因此，本文采以“一元二面多维多层次”的人体系统模型为核心，来综合集成相关人体在结构或功能、身体或情志等方面的症状，由此构建了《健康指数问卷》。

具体来说，问卷的条目主要是参考了“十三五”规划教材《中医诊断学》一书，以及其他经典量表中的条目，并且经过小样本测试、中西医专家咨询等过程，多次迭代，最终纳入了 106 个常见的人体症状条目，并且根据症状的严重程度，划分为五个严重等级。

实际使用时，受测者根据近一个月的实际情况，对该问卷中每条条目的客观症状出现的频率进行选择评分。0 分是几乎从不出现，1 分是偶尔出现，2 分是有时会出现，3 分是经常出现，4 分为总是出现。

回答结束后,对问卷条目的选择结果 $x_i$ 与该条目的严重程度 $I_i$ 的加权加和 $\sum_1^N(x_i * I_i)$ ,就代表了该人士偏离理想健康状态的值。问卷的总分减去该值,则代表了人体的健康情况。为了方便人们使用和理解,将这个数值归一化到百分制,即有:

$$H = 100 * \frac{S - \sum_1^N(x_i * I_i)}{S} \quad (6.1)$$

其中,  $x_i$ 是每道题目的实际得分,  $I_i \in [1,2,3,4,5]$ 为每道题目的严重等级,  $N = 106$ 是问卷条目的总数,  $S = 2120$ 是问卷的满分。

本文将根据 6.1 式计算出的数值定义为“健康指数H”,它反映了人体系统距离理想健康状态的情况。当受测者没有任何症状时,健康指数 $H = 100$ ,为理想的健康状态;而如果受测者出现了所有症状,并且频率都属于“总是出现”时,健康指数 $H = 0$ ,对应的是离理想健康状态最远的情况。

特别的,针对每日测量脉搏波的实验,由于受测者都是健康的青年,在综合考虑测量的方便性和可持续性的基础上,我们从《健康指数问卷》的条目中选取了 57 条青年可能会出现的症状,形成了《健康指数问卷(日常测量版)》。同样,常测人员在每日测量完成脉搏波实验后填写问卷,并且根据 6.1 式就可以计算出日变健康指数,此时 $I_i = 1$ ,  $N = 57$ ,  $S = 171$ 。

《健康指数问卷》的特点是以人体复杂系统模型为核心来综合集成相关症状。各条目涉及的症状都是客观的,反映的是与理想健康状态的偏离。由此就形成了对于个体多维多层次的刻画,据此计算的人体健康指数H则实现了对人体偏离理想健康态的刻画。北京大学医学部刘奕君博士<sup>[164]</sup>以同样的方法研发了作业和谐问卷,并将此问卷用以刻画 734 名大学生的健康状态,与国际流行量表 DASS-21 相比,结果具有很好的一致性(见图 6.1),并充分证明了问卷制作的有效性。

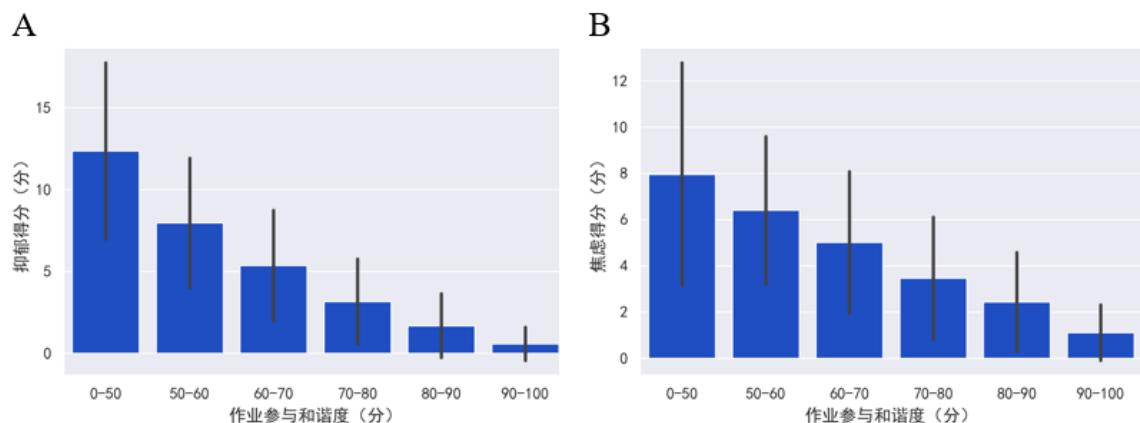


图 6.1 以同样方法研制的问卷结果与国际流行的 DASS-21 量表结果有很好的一致性<sup>[164]</sup>

## 6.2 刻画群体的稳定健康指数

本节我们使用脉搏波的序参量模型，来对群体的长期稳定的健康指数进行刻画。

### 6.2.1 数据情况介绍

首先对本节使用的数据进行简要的介绍。在 2021 年 5 月-6 月，我们在社区随机采取了 23 位受测者的脉搏波数据。并且受测者在实验员的帮助下，根据近一个月来的身体情况，填写了健康指数问卷数据。表 6.1 展示了这 23 位受测者的性别、年龄和 BMI 等基本信息。

表 6.1 群体测量人员及其数据情况

	性别			BMI			年龄		
	男	女		$\leq 18.4$	18.5-23.9	$\geq 24$	18-22	23-30	$\geq 30$
样本量	7	16		4	14	4	9	7	6

注：缺少 1 样本的 BMI 和年龄数据。

对于上述数据，根据 5.4 节介绍的数据预处理方法，处理脉搏波数据，并且根据 6.1 式计算出个体的健康指数。

### 6.2.2 建立序参量模型

根据上一章介绍的脉搏波序参量模型的建模步骤，首先需要根据待研究的问题，选取合适的序参量，建立序参量相空间。这里由于刻画的是长期的比较稳定的健康状态，所以我们选取了六部平均后的脉位  $\bar{P}$  和脉力  $\bar{F}$ ，来建立序参量相空间：

$$\chi = \{\bar{P}, \bar{F}\} \quad (6.2)$$

在这个相空间中，运用 K-means 聚类方法将所有样本数据自动聚为 3 类，如图 6.2 所示。同一种颜色的数据点，表示属于同一种亚类。进一步，根据脉搏波的序参量模型（5.7 式-5.10 式），确定出每一类的理想点坐标（图 6.2a 中的五角星），以及线性率的参数（见表 6.2）。

### 6.2.3 模型的验证

首先，由图 6.2a 可以看出，在序参量相空间中自动聚类得到的三个亚类，具有明显的样本聚集特征。如果以所有样本的均值画两条参考线（图 6.2a 中黑色虚线），则各亚类处于相空间的不同位置。由于这里使用的序参量（脉位  $\bar{P}$  和脉力  $\bar{F}$ ）的物理含义十分明确，所以我们可以直接根据样本所处的位置，来解释各亚类所反映的其脉搏波运动的特点。

例如，亚类 1（图 6.2a 中的紫色圆圈）处于相空间的左下角，代表其脉搏波的主要特点是“脉位浮”和“脉力弱”，对应于中医脉诊理论中的“浮弱脉”。而亚类 2（图 6.2a 中的红色圆圈）恰好相反，它代表的脉搏波的主要特点是“脉位沉”和“脉力强”，对应于中医脉诊理论中的“沉强脉”。在中医脉诊理论中，这些脉象具有明确的临床诊断含义，可以直接供临床使用。

同时，我们可以看到，相空间中的左上角是空缺的，而它所代表的脉搏波的主要特点是“脉位浮”和“脉力强”，对应于中医脉诊理论中的“浮强脉”。根据中医理论，这类脉象主要常见于刚有外感病且体质较为强壮的人，这也就解释了我们采集的样本中没有这类脉象的原因。

进一步，根据模型确定出每个亚类的理想点（图 6.2a 中的五角星），以及据此计算出的各亚类中样本与理想点之间的脉搏波偏离度  $d_i$ 。图 6.2b-f 的结果显示，所有亚类的样本的脉搏波偏离度  $d_i$  与健康指数  $H_i$  之间，都具有很好的负线性关系 ( $R^2 > 0.93$ )。即脉搏波偏离度  $d_i$  越大，则健康指数  $H_i$  越低。而脉搏波偏离度  $d_i$  越小，则健康指数  $H_i$  越高。这一结果有力地证明了模型的假设。

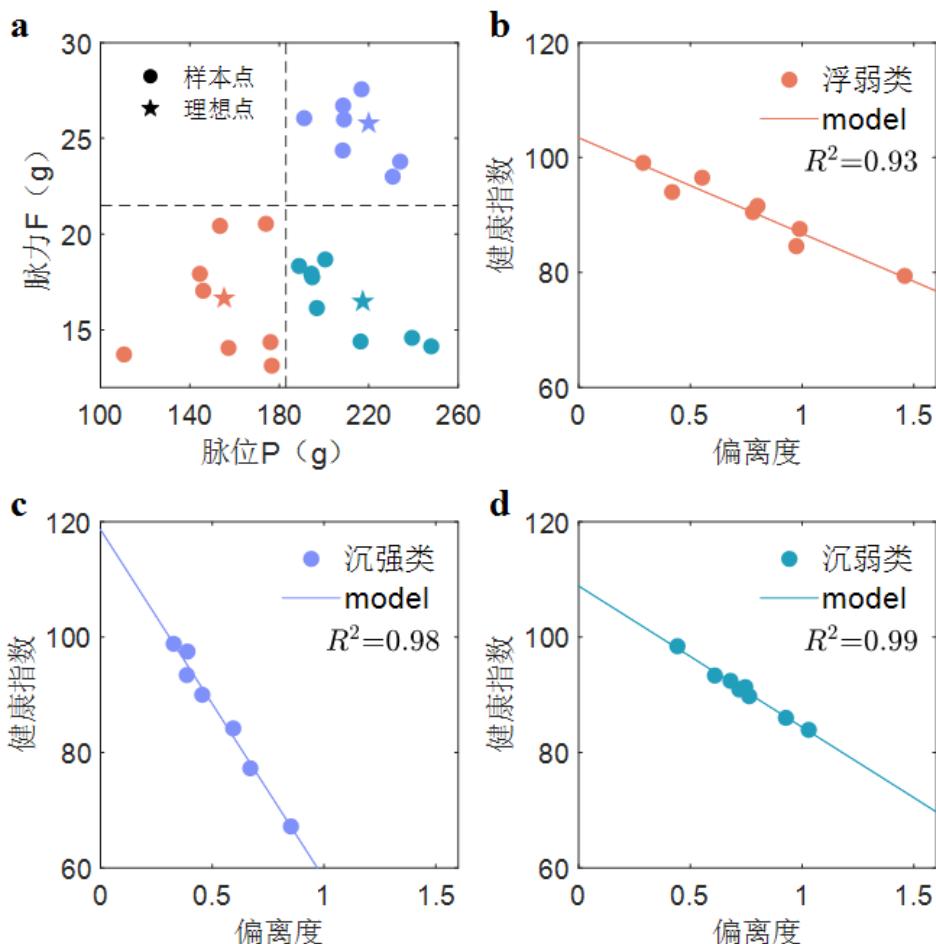


图 6.2 脉搏波序参量模型对群体的长期稳定的健康指数进行刻画的结果 ( $N=23$ )。图中圆圈代表实测数据值，五角星为计算出的理想点位置，不同颜色代表所属于不同的亚类。直线代表模型的拟合结果。**a.** 所有样本的脉搏波序参量值在相空间中的分布结果。**b-d** 分别为各亚类样本的脉搏波偏离度与健康指数之间的线性关系。其中，脉位为脉搏波的基线压力 (**g**)，脉力位主波振幅 (**g**)。**e.** 为设置不同聚类簇数时的拟合误差。

需要指出的是，上述结果的获得，是因为根据“序参量相空间”划分了亚类，并且确定出了每一个亚类的理想点的位置。这说明，理想的健康水平，并不是一个固定的状态，而是在每一类的人群中，都存在一个较为理想的健康态。所以，不同个体之间的健康状态的比较，应该在同一个亚类的样本中进行比较，而不应该是不进行区分的直接进行“大数平均”。

表 6.2. 不同亚类的样本量和模型参数

不同亚类	Cluster 1	Cluster 2	Cluster 3	所有样本
样本量	8	7	8	23
敏感因子 $k$	-16.67	-60.48	-24.48	-33.87
理想健康指数 $H_{max}$	103.5	118.6	108.9	—
拟合优度 $R^2$	0.93	0.98	0.99	0.97

#### 6.2.4 与传统分类结果相比较

无论是基于人口学特征（如年龄、性别），还是流行的健康指标（如 BMI），这些对人群进行分类的传统方法经常被用于人体健康的相关研究中。所以，我们也通过年龄、性别和 BMI 对上述样本进行分类（见表 6.1），然后用同样的方法（式 5.7-5.10）刻画个体的健康指数，由此来对比不同分类方法的特点。

结果显示，使用了序参量相空间分类方法的刻画偏差，在选取不同的分类数目时，其平均相对误差（MRE）都显著地低于使用传统分类方法的刻画偏差（见图 6.3）。这说明，相对于年龄、性别和 BMI，基于序参量相空间的分类方法能够更加反映同一类群体在健康方面的共同特征。所以，基于序参量相空间的分类方法，本文才能够发现脉搏波偏离度与健康指数之间的普适线性率（见图 6.2）。

同时，这一结果还说明了类别的数目并不是影响刻画健康指数的关键因素。虽然似乎类别的数目越多，每一类的样本就越少，更容易被线性模型所刻画。但是，基于序参量相空间的分类结果，都要显著低于传统分类的结果。这说明影响刻画结果的关键因素不是分类的数目，而是用什么量去做分类。

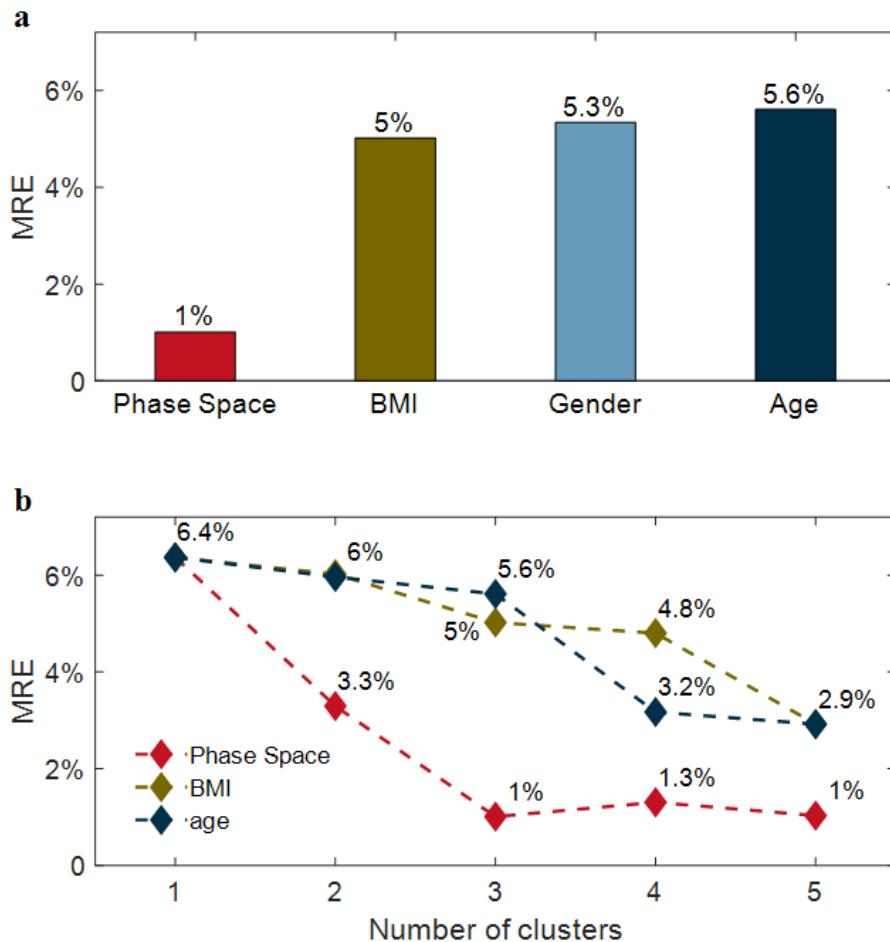


图 6.3 不同分类方式和分类数目的预测误差的比较 ( $N=23$ )。a. 不同分类方式刻画健康指数的平均相对误差 (MRE)。b. 不同分类数目刻画健康指数的平均相对误差 (MSE)。

### 6.2.5 各亚类的健康态特点

进一步，根据受测者所填问卷情况，可对 3 个亚类人群的症状分布进行探讨。根据症状在三个亚类出现频率的高低和差异，我们挑选出了 7 个特异性症状。图 6.4 展示了三个亚类 7 个症状的频率谱。可以看出，三个亚类的频率谱明显不同。对于亚类 1 和亚类 2，除了“突然站立头晕”这个症状，亚类 1 频率较高的症状，在亚类 2 中频率都较低。对于亚类 1 和亚类 3，除了“心烦”这个症状，在亚类 3 中出现频率较高的症状，在亚类 1 中都较低。对于亚类 2 和亚类 3，“心烦”、“突然站立头晕”、“精神萎靡”的症状频率明显不同。

进一步，我们可以总结出三个亚类的特异性气血特点。表 6.3 总结了三个亚类的高频和低频特异性症状，其中高频特异性症状指该症状在该亚类中频率高，在另外两个亚类频率低；低频症状指该症状在该亚类频率低，在另外两个亚类中频率高。如症状

“突然站立头晕”为沉弱类人群的特异性高频症状，这说明该症状在亚类三中频率高，在另外两个亚类中频率都较低。

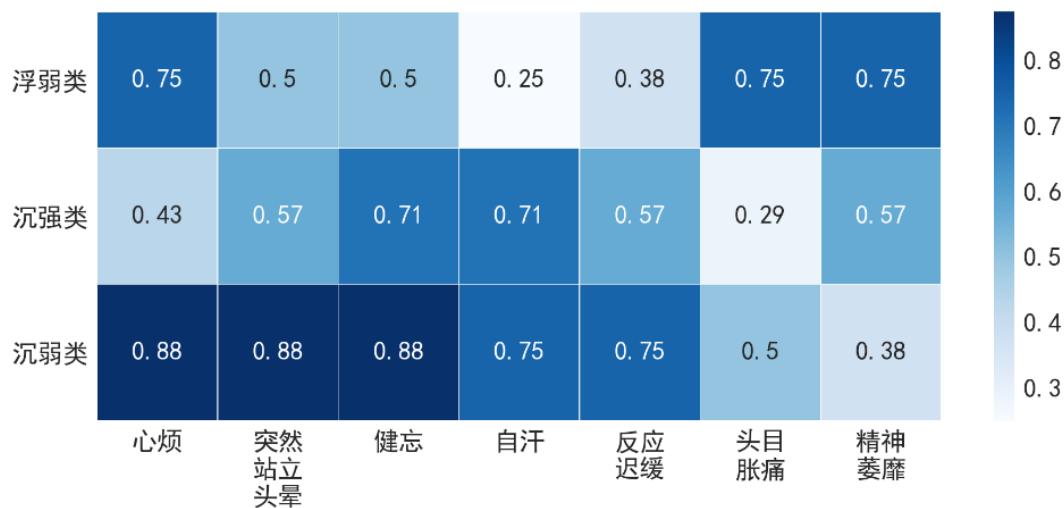


图 6.4 各亚类的症状谱分布。每一行分别为三个亚类的症状分布谱，格子中的数字和颜色深浅代表了相应症状出现的频率大小。

到此，我们总结出了各亚类的特异性症状，这些症状从一个角度说明了不同亚类人群的具有不同的健康状态，同一亚类人群具有相似的健康状态。亚类的部分特异性症状与该亚类脉搏特点具有一致性，如沉弱类的亚类特异性症状为“突然站立头晕”，根据中医理论，该症状往往是由气血郁闭或者虚损，不能上荣头部导致，这与沉弱类脉搏表征气机郁闭在内且虚弱的特点是一致的。再比如，沉强类没有特异性的高频症状，根据中医理论，沉强类脉搏特点脉位偏沉、脉力强，正常情况下代表人体气机收敛状态好，气机力量强。未来需对沉强类人群进一步进行亚类划分并讨论，更深入探讨个体的健康特点。上述结果为未来对脉搏如何与健康状态的对应提供了方向，未来可在更大的数据集上进行讨论验证。

表 6.3 不同亚类的样本量和模型参数

亚类编号	特异性症状	
	高频	低频
亚类 1（偏浮弱类）	头目胀痛	自汗、健忘、反应迟缓
亚类 2（偏沉强类）	健忘、自汗	心烦
亚类 3（偏沉弱类）	突然站立头晕	精神萎靡

### 6.3 刻画个体的日变健康指数

在上一节成功刻画了群体的长期稳定的健康指数的基础上，我们进一步使用脉搏波序参量模型对某位日常测量的实验员在长达 4 个月内连续测量的每日“健康指数”进行了刻画。

#### 6.3.1 数据情况介绍

本节使用的数据是一位青年日常测量人员在 2019 年 11 月—2020 年 3 月份进行连续测的数据集。该名测量人员一般固定在每天晚上 21:00-23:00 之间测量脉搏波（详见 5.4 节），然后填写《健康指数问卷（日常测量版）》。具体受测者和测量数据的情况，如表 6.4 所示。

表 6.4 日常测量人员及其数据测量情况

性别	年龄	BMI	持续时间（天）	测量时间（天）	测量占比
女	27	21	117	90	77%

对于上述数据，根据 5.4 节介绍的数据预处理方法，处理脉搏波数据，提取了脉搏波的基线压力 (g)、主波振幅 (g) 和波动周期 (s)。数据见附录 B。并且根据 6.1 式计算出个体每日的健康指数。

#### 6.3.2 构建序参量模型

根据 5.2 节介绍，传统脉搏波研究主要进行的是二分类研究，而定量刻画每日健康状态的波动，则是一件充满挑战性的事情。这是因为相比于不同个体之间的健康指数的差别，个体的每天健康变化是一个“小量”。这个“小量”被淹没在脉搏波每天的较大的涨落之中。

为此，我们仍然是按照脉搏波序参量建模的步骤，首先构建了包含六部脉搏波的主导序参量（脉位 P、脉力 F、周期 T）在内的序参量相空间：

$$\chi = [P_1, P_2, P_3, P_4, P_5, P_6, F_1, F_2, F_3, F_4, F_5, F_6, T] \quad (6.3)$$

其中， $P_i$  代表脉位， $F_i$  代表脉力。 $i = 1, 2, 3, 4, 5, 6$ ，依次代表左手、右手的寸关尺六个测量部位。考虑到在测量期间脉率不会发生较大的变化，所以这里的脉率  $T$  为经过六部平均后的脉率。

首先，我们使用了经典的无监督聚类方法 K-means 在序参量空间中对所有样本的脉搏波数据进行聚类，共形成 5 个亚类。为了更好的展示不同亚类的分布特点，我们使用了常用的降维可视化方法 t-SNE 对数据在高维的序参量空间  $\chi$  的分布情况进行了降维展示。如图 6.5 所示，同一种颜色的数据点，表示属于同一种亚类。可以看到，5

个亚类之间的距离比较接近，但都能够比较清楚地区分开。除了第 3 个亚类（见图 6.5 中的红色点）距离中心较远外，其他 4 个亚类都是紧密围绕着中心分布在不同的方位上。这说明个体在序参量相空间的位置是相对比较其中的，每日测量的数据都是围绕着“常态”而偏离，验证了第五章提出的假设。

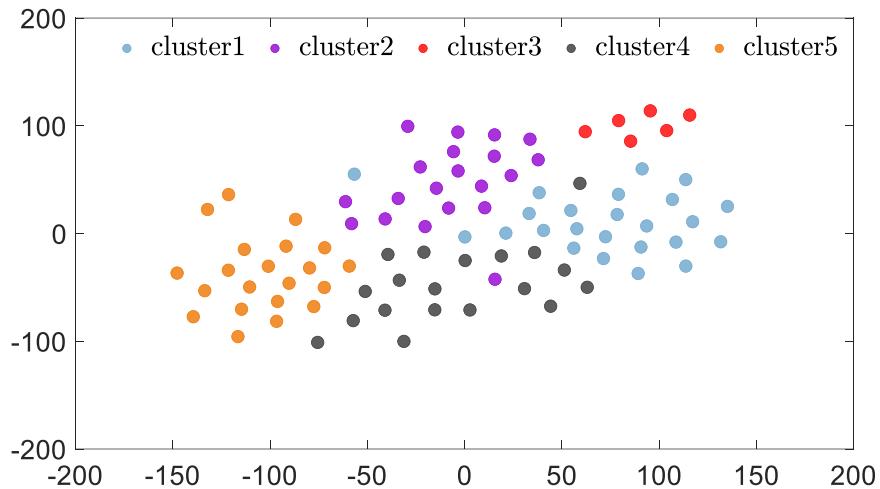


图 6.5 连续测量的脉搏波数据经过 t-SNE 降维后的可视化结果 ( $N=90$ )。

### 6.3.3 普适的线性率

根据脉搏波序参量模型 (5.7 式-5.10 式)，可以计算出每个亚类的理想点  $I$  的位置，以及模型参数（见表 6.5）。如图 6.6 所示，5 个亚类的所有样本的脉搏波偏离度  $d_i$  与其健康指数  $H_i$  之间都具有普适的线性关系。除了第 1 类的拟合优度  $R^2$  是 0.5 以外，其余 4 类的拟合优度  $R^2$  均大于 0.85。这一结果表明，通过测量脉搏波，序参量模型就可以精确地刻画出每天健康指数  $H$  的波动，这有力地验证了序参量模型的有效性。

进一步，我们把脉搏波序参量模型对于健康指数  $H$  的刻画结果，展示在了时间轴上，如图 6.7 所示。首先，该受测者的健康指数的平均分为 94.9 分，标准差为 1.8 分，说明该受测者的健康状况良好且比较稳定。然而，如果从时间演化的角度来看，个体的健康指数随着时间变化存在着连续变化的小尺度结构，以及相邻两天突然剧变的复杂行为。对于这种存在着脉动较大的复杂演化行为，脉搏波序参量模型在长达 4 个月的数据测量时间范围中，模型刻画的相对误差基本都在 2% 以内，平均相对误差仅为 0.6%。这显示了脉搏波序参量模型的强大的刻画能力。

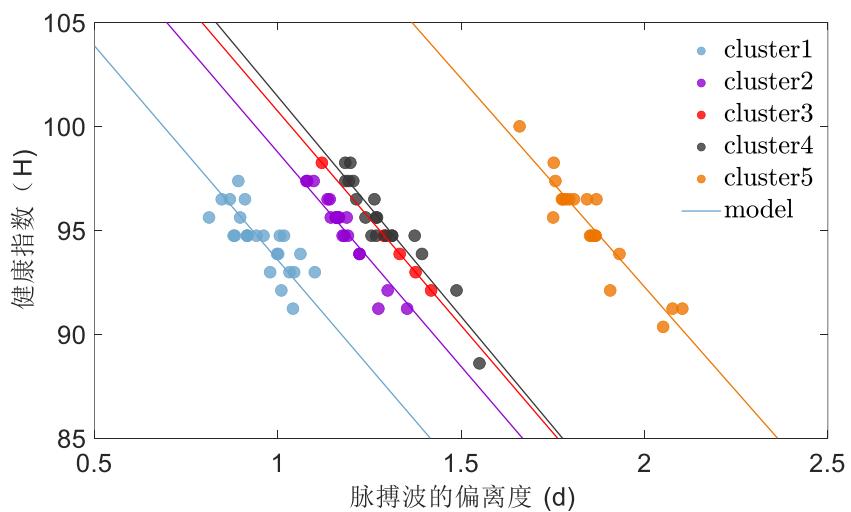


图 6.6 各亚类样本的脉搏波偏离度与健康指数之间呈线性关系 ( $N=90$ )。

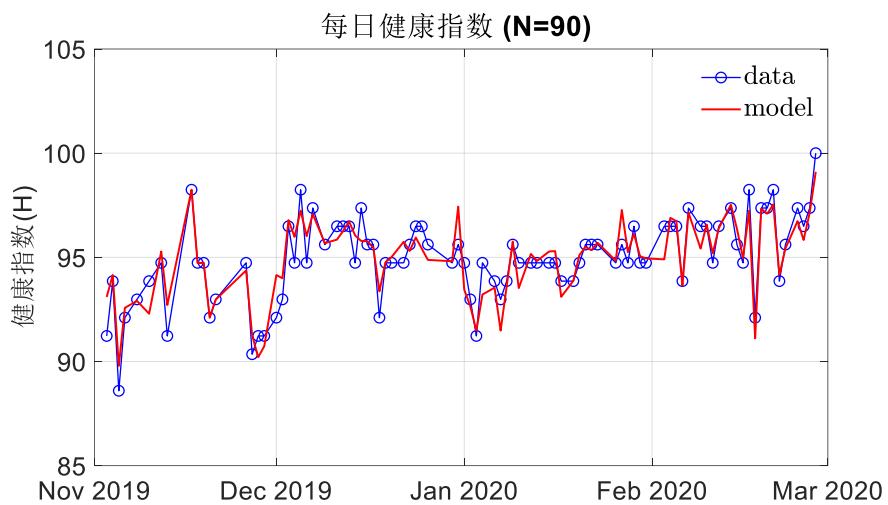


图 6.7 脉搏波序参量模型对个体日变健康指数的刻画结果 ( $N=90$ )。图中横坐标为测量日期，纵坐标为日变健康指数。蓝色线为实际测量结果，红色线为模型刻画结果。

#### 6.3.4 各亚类的关系

我们还发现这 5 个亚类所呈现的线性关系的斜率  $k$  都十分接近，约为 -20（见表 6.5）。根据模型定义（见 5.9 式）， $k$  为反映脉搏波偏离度变化与健康指数变化的敏感因子。然而，数据显示，所有的亚类敏感因子  $k$  竟然都是一致的，这预示着敏感因子  $k$  是所有亚类所共同满足的一个普适关系。由于  $k$  是一个联系局部的脉搏波测量结果和整体的健康指数  $H$  之间的一个关键变量，值得在未来通过更多的数据样本进行检验。

表 6.5 不同亚类的模型参数和数据情况

不同亚类	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	所有样本
样本量	24	20	6	19	21	90
敏感因子 $k$	-20.6*	-20.6*	-20.6	-21.1	-20.1	-20.1
平均理想健康指数 $\bar{H}$	94.5	95.0	94.4	95.3	95.2	94.9
拟合优度 $R^2$	0.50	0.90	0.99	0.89	0.87	0.83

\* 第 1 类和第 2 类的  $k$  值为其他三类的  $k$  值的平均值

另一个模型参数是线性率在  $y$  轴上的截距  $H_{max}$ 。根据模型定义（见 5.9 式），该截距代表了该亚类的理想点的最高健康水平，即偏移度为 0 的时候的健康指数  $H_{max}$ 。数据显示，不同亚类的理想点的健康指数值  $H_{max}$ ，与该亚类所有样本的脉搏波偏移度的平均值  $\bar{d}$ ，符合很好的线性关系（见图 6.8）。

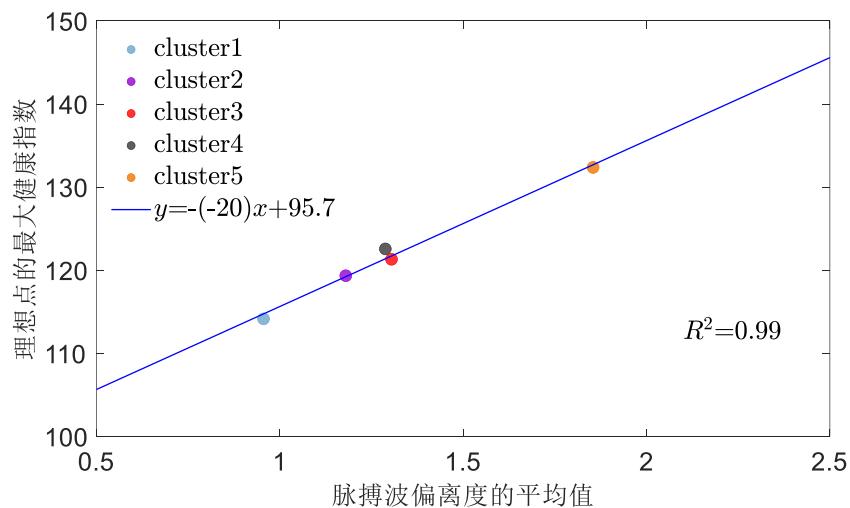


图 6.8 各亚类的脉搏波偏移度的平均值与理想点的最大健康指数之间呈线性关系。

这一结果意味着，当该亚类的所有样本的脉搏波偏移度的平均值  $\bar{d}$  越大时，即数据分布距离理想点越远时，该亚类的理想点的健康指数  $H_{max}$  就越大。因此，我们定义了重新尺度化的距离  $d_i - \bar{d}_i$  和健康指数  $H_i - \bar{H}_i$ ，从而能够使得 5 个亚类的数据都重叠到一条直线上，如图 6.9 所示。这一结果告诉我们，不同亚类虽然有着不同的理想点，并且由此计算的簇内距离  $d_i$  不一样，但是它们都遵循一个更为普遍的线性律。普遍线性率的斜率，即敏感因子  $k$ ，是一个重要的表征个体特征的参数，而重新尺度化的距离  $d_i - \bar{d}_i$  是脉搏波的一个重要的特征距离。

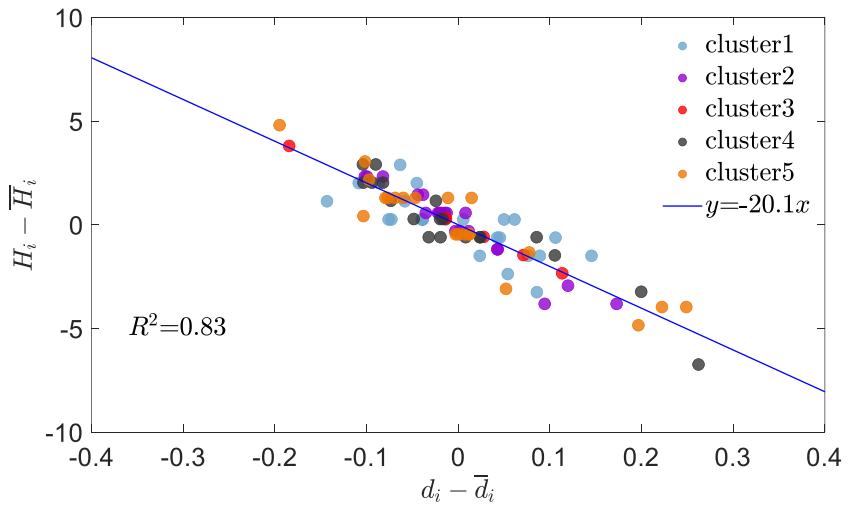


图 6.9 重新尺度化后所有样本的脉搏波偏离度与健康指数之间呈线性关系 (N=90)。

### 6.3.5 特征距离的时间演化

根据上述讨论,本研究总结发现了脉搏波在序参量相空间中的运动存在着两类特征距离,分别反映了重要的人体健康情况的信息。

首先,对于所有测量数据(N=90)而言,在序参量空间中,每次脉搏波测量的样本 $S_i$ ,与所有样本的中心 $C$ 之间的偏离度 $d_{center}$ 之间存在着明显的时间周期演化特征(见图6.10a)。我们用一个简单的傅里叶级数函数对这一演化进行了定量地刻画(见图6.10a中的黑线),显示其周期T约为35天。所有样本的中心 $C$ 可以看作是该个体在这段时期内的“常态”,而每天脉搏波测量的结果都是围绕着在这一“常态”附近偏离。而这种偏离具有一种随时间周期运动的“基态”。特别地,当偏离较大而远离周期运动时,如图6.10中的红色数据点,正好对应着一个独特的亚类。

其次,对于各亚类而言,每个亚类的样本点 $S_i$ 与该亚类的理想点 $I$ 之间的偏离度 $d_i$ ,也存在着明显的时间演化特征(见图6.10b)。我们仍然用一个简单的傅里叶级数函数对这其进行定量地刻画(见图6.10b中的黑线),显示其周期T约为70天,恰好为中心偏离度 $d_{center}$ 的周期的2倍。更重要的是,这一演化行为,能够清楚地将不同的亚类进行区分。例如,第5类(见图6.10b中的黄色点)主要出现在波峰附近,第3类(见图6.10b中的红色点)主要出现在上升沿,而其他类则主要出现在较为平缓的部分。

由此可见,不同亚类具有不同的演化行为,脉搏波序参量模型正是因为抓住了各亚类演化的各自特征,所以才能将日变健康指数进行良好地刻画。对于这些特征距离,值得在今后的研究中利用更丰富的样本来探究人体脉搏波运动的规律。

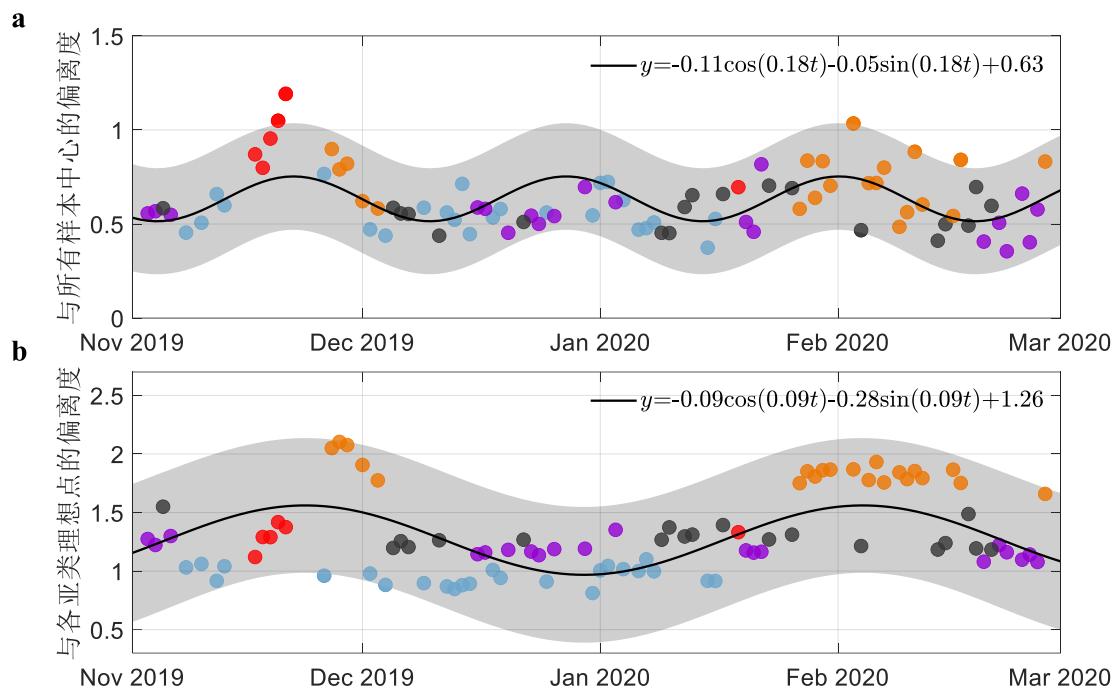


图 6.10 特征距离的演化特征。图中横坐标是时间，纵坐标是特征距离。实心圆圈代表实际脉搏波测量的数据值，不同颜色代表所属于不同的亚类。黑色曲线是傅里叶级数的刻画结果，灰色条带是 95% 的置信区间。

### 6.3.6 日变健康指数的预测

进一步，我们采用了 k 折交叉验证（k-fold cross validation）的方法，对模型的可靠性和稳定性进行了验证。将数据随机平均分成 18 份，分别对其中的 17 份进行模型训练，剩下的 1 份 ( $N=5$ ) 进行验证。结果显示，模型刻画的平均相对偏差为  $1.9\% \pm 1.7\%$ ，并且模型参数  $k$  的平均值为  $-20.5 \pm 0.5$ ，这有力地支持了模型的有效性和可靠性。

进一步，我们采用 2022 年 1 月 4 日-22 日之间测量的 16 条数据进行了模型预测，同时与经典的机器学习方法——支持向量机（SVM）的结果进行了比较（见表 6.6）。结果显示，脉搏波序参量模型预测结果的平均相对误差和平均均方差都比支持向量机的偏差小，显示了更高的预测精度。

表 6.6 模型预测的结果比较

	Mean Relative Error	Mean Square Error
脉搏波序参量模型	4.1%	20.82
支持向量机	4.6%	24.73

除了预测精度高以外，更重要的是，脉搏波序参量模型由于建模过程十分清楚，模型参数简单且都具有明确的物理意义，所以相比传统的机器学习方法，能够提供更丰富的信息，从而帮助我们对脉搏波信号开展不断深入的分析，加深对人体信号的理解。

## 6.4 应用于分析其他人体波动信号

进一步，我们将上述脉搏波序参量模型的建模方法（见 5.5 节），应用于分析其他人体生理波动信号，即人体手指的气体放电可视化（Gas Discharge Visualization, GDV）信号，也取得了很好的效果。

### 6.4.1 数据情况介绍

气体放电可视化（Gas Discharge Visualization, GDV）技术是一种光电成像系统（见图 6.11）。它的主要工作原理是，通过在手指与绝缘板之间加一个瞬时高压，将二者之间的空气击穿而放电成像。作为一种新兴的健康测评技术，目前被应用于识别高血压、糖尿病、冠心病、肿瘤等疾病。

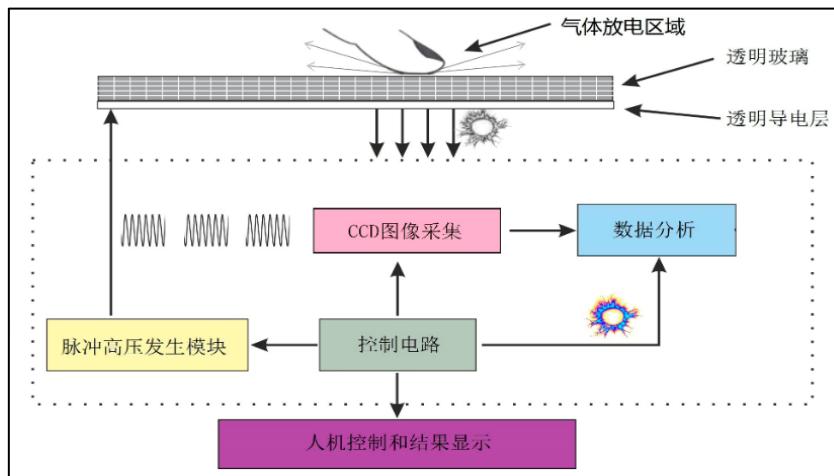


图 6.11 GDV 原理示意图<sup>[165]</sup>

2021 年 5 月-6 月，我们在社区随机采集了不同年龄段人群的左右手的 GDV 原始图像，如图 6.12 所示。图中黑色为背景颜色，十个蓝色光晕分别为手指附近的气体放电图像。

在 GDV 测量结束后，受测者在实验员的帮助下，根据近一个月内的身体情况，填写《健康指数问卷》（见 6.1 节）。共计得到 20 个有效样本，其性别、年龄及 BMI 的分布情况如表 6.7 所示。

对于 GDV 原始图像，本文通过图像处理，得到了原始图像的十指辉光的平均亮度

$\bar{I}$ 和十指平均面积 $\bar{S}$ , 用于建模分析。由于这部分内容不是本文的重点, 为了行文更加通畅, 具体的图像处理过程请见附录 A。

对于健康指数问卷, 本文通过 6.1 式计算了每例样本的健康指数值。

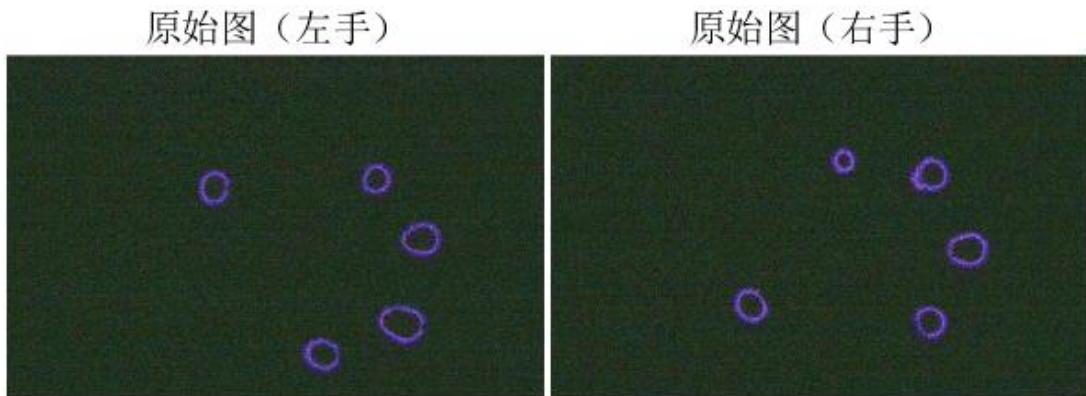


图 6.12 GDV 信号的原始图像

表 6.7 同时测量脉搏和 GDV 的样本数据情况

	性别		BMI			年齡		
	男	女	$\leq 18$	18-23.9	$> 23.9$	18-20	20-25	$> 25$
样本量	8	12	5	8	7	6	9	5

#### 6.4.2 建立序参量模型

跟脉搏波序参量模型的建模过程一样, 首先需要选取合适的序参量, 建立序参量的相空间。这里我们选取了 GDV 原始数据的最直观的两个量: 十指辉光平均面积 $\bar{S}$ 和十指辉光平均亮度 $\bar{I}$ , 来建立序参量相空间:

$$\chi = \{\bar{S}, \bar{I}\} \quad (6.4)$$

在这个相空间中, 运用 K-means 聚类方法将所有样本聚为 3 类, 如图 6.13 所示。同一种颜色的数据点, 表示属于同一种亚类。进一步, 根据序参量模型的计算方法 (5.7 式-5.10 式), 确定出每一类的理想点坐标 (见图 6.13a 中的五角星), 以及线性率的参数 (见表 6.8)。

### 6.4.3 普适的线性率

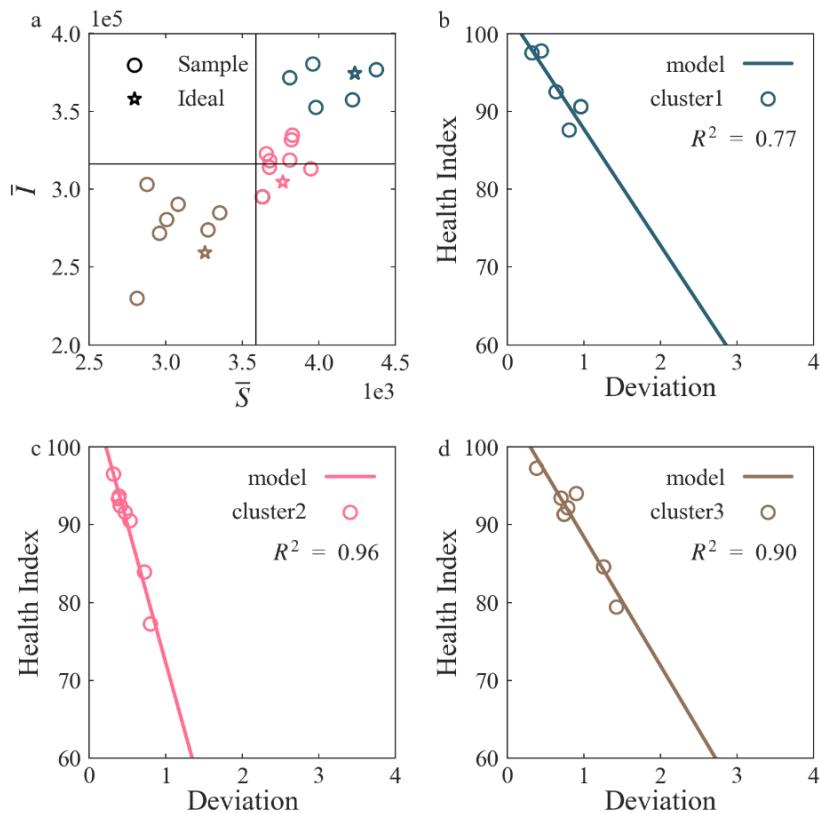


图 6.13 对于 GDV 信号构建序参量模型的刻画结果。图中圆圈代表 GDV 数据，五角星为计算出的理想点位置，不同颜色代表不同的亚类。直线是模型的刻画结果。a. 各亚类 GDV 数据在相空间的分布结果。b-d 分别为三个亚类的偏离度与健康指数之间的线性关系。

表 6.8 不同亚类的模型参数和数据情况

不同亚类	Cluster 1	Cluster 2	Cluster 3
样本量	5	8	7
敏感因子 $k$	-14.9	-35.5	-16.5
平均理想健康指数 $\bar{H}$	102.7	107.7	105.0
拟合优度 $R^2$	0.77	0.96	0.90

由图 6.13a 可以看出，在序参量相空间中自动聚类得到的三个亚类，具有明显的聚集特征。以所有样本的均值画两条参考线（图 6.13a 中黑色虚线），可以看出三个亚类主要分布于一三象限，即样本的十指辉光平均面积  $\bar{S}$  与十指辉光平均亮度  $\bar{I}$  的变化趋势基本一致。与脉搏波的结果（见图 6.2）类似，GDV 序参量模型计算出的 GDV 偏离度  $d^{GDV}$  与其稳定健康指数  $H^{GDV}$  之间，也具有普适的线性率关系 ( $R^2 > 0.80$ )。并且，三个亚类的敏感因子  $k$  的数值十分接近。

脉搏波信号和 GDV 信号，属于基于不同测量原理的人体生理波动信号，但都反映了人体健康状态波动。通过序参量模型，本文都发现了信号偏离度与健康指数之间的普适线性关系。这预示着，人体生理波动信号在“人体序参量相空间”中隐含着高度结构化的特征，值得在未来通过更丰富的样本数据来验证。

与脉搏波的结果类似，使用了序参量相空间分类方法的刻画偏差，都显著地低于使用传统分类方法的刻画偏差（见图 6.14）。这说明对于序参量相空间的分类方法具有良好的泛化能力，在 GDV 这一测量工具所得到的健康数据上依然适用。

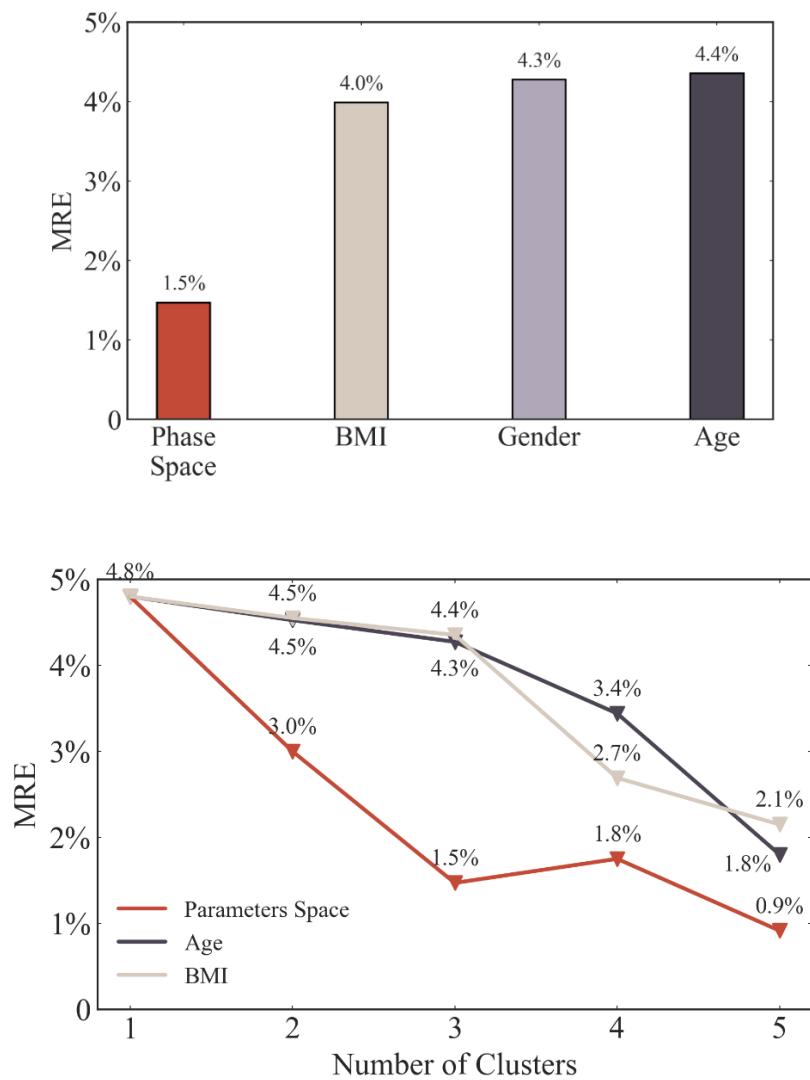


图 6.14 不同分类方式和分类书库的预测误差的比较 ( $N=20$ )。a. 不同分类方式刻画健康指数的平均相对误差 (MRE)。b. 不同分类数目刻画健康指数的平均相对误差 (MRE)。

#### 6.4.4 不同亚类的特点

根据 GDV 受测者所填问卷情况，我们对 3 个亚类人群的症状分布进行探讨。根据

症状在三个亚类出现频率的高低和差异，我们挑选出了6个特异性症状，如图6.15所示。可以很直观的发现，明亮饱满类的高频（频率>0.8）症状主要为“自汗”、“尿频”，适中类的高频症状为“心烦”和“注意力下降”，暗弱稀疏类在六个症状上都具有较高的频率。当症状在某一类呈现高频特征的时候，在其他的类别中该症状的频率有明显的降低，这说明每个亚类都具有特异的高频症状，互相之间的区分度较高。GDV亚类在症状谱上也具有明显的区别，反映了使用序参量相空间所划分的亚类能够反映不同类型的健康特点。这一套建模和分析的方法，有助于开展个性化的健康辨析，有望未来通过更多的数据来进行验证和优化。

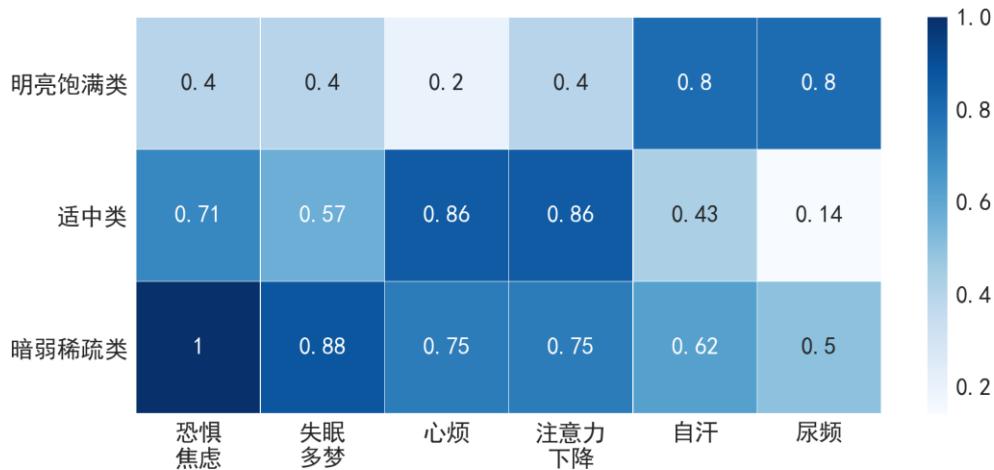


图 6.15 各亚类的症状谱分布。每一行分别为三个亚类的症状分布谱，格子中的数字和颜色深浅代表了相应症状出现的频率大小。

## 6.5 本章小结

基于脉搏波刻画人体健康状态的细致变化，是一件充满挑战性的事情。一方面，由于人体是一个高度开放的复杂巨系统，脉搏波本身很容易受到自身状态和外界环境所带来的影响而发生较大的涨落。另一方面，脉搏波信号本身具有一定的涨落，复杂度很高。已有的研究，主要是基于大数统计平均的方法，来开展“二分类”（健康与病态）的研究，尚难以对人体健康数据开展连续的定量刻画（见5.2节）。面对这一挑战，我们主要开展了以下几方面的工作：

首先，我们系统地设计了实验方案，来获得较高质量的数据集。除了常见的群体测量实验外，我们还开展了日常测量实验，即实验人员每天定时定点地开展脉搏波测量，尽可能地减少由于测量条件变化所带来的影响。本文也是第一次公开报告了个体连续一年的测量结果，这也是本研究的一大特色。另一方面，针对传统基于生理检测指标对人体健康状态标记具有的碎片化等局限性，本研究团队基于人体复杂系统模型专门设

计了一套由客观症状所组成的健康指数问卷。由此保证了能够抓住个体每日健康波动的“小量”。

其次，我们将人体序参量理论应用于脉搏波分析，建立了脉搏波序参量模型。本章的结果显示，脉搏波序参量模型能够精确地刻画不同个体的常态健康指数，以及某个个体长期连续变化的健康指数。更重要的是，相比传统的基于大数统计平均所确定的形式复杂、意义不清的参数，本文只用到了脉搏波的“基线压力”“主波振幅”和“搏动周期”这三个最基本的参数。由于这三个参数的生理含义非常明确，所以通过模型刻画结果可以对脉搏波信号开展不断深入的理解。

最后，通过对比不同分类方法的结果，我们的模型还发现，基于序参量相空间的分类方法，更能够抓住反映人体健康水平的重要特征。无论是对群体的脉搏波数据，还是个体长期的脉搏波数据，乃至与其它来源的 GDV 信号，本文都发现在序参量相空间中，信号的偏离度与健康指数之间存在普适的线性率。这一结果证实了模型的有效性和普适性，揭示了人体生理波动信号中隐含的高度结构化的特征。这为开展融入医学经验和知识的大规模机器学习提供了理论基础。

本研究同样存在一些不足。首先，由于疫情的影响，本文在开展群体实验测量时受到了很大的挑战，所以群体研究纳入的样本量较小。因此本文重点是针对个体长期测量的数据开展了刻画。其次，我们自主设计的健康指数问卷在将来还需要通过大样本测试，来证实健康指数问卷的信度和效度。



## 第七章 总结和展望

本文针对“人体复杂系统”开展了两个方面的研究。其一，针对新冠疫情传播系统和死亡率演化系统，讨论了个体是如何受外界环境作用而演化。其二，针对人体脉搏波系统，讨论了个体脉搏波的运动特点以及如何反映人体宏观的健康状态。通过应用于这两个跨尺度的系统，探索人体序参量理论的普适意义。本章对上述研究内容和成果进行了梳理和总结，并对未来的研究工作进行了计划。

### 7.1 主要结论

将人体序参量理论应用于新冠疫情传播系统，本文主要获得了以下结论：

(1) 本文发现并刻画了疫情中晚期由国家干预措施所主导的新的对称破缺过程，提炼出国家干预过程的三个关键序参量：隔离度 ( $G$ )，治愈度 ( $H$ ) 和抢救度 ( $R$ )，建立了远比传统传染病模型更为精确的 GHR 传染病动力学模型，得到疫情数据的精确支持。在将其应用于国内外疫情演化的精细对比中，揭示了新冠疫情传播中的“序参量-转换速率-群体数量”三个层次相互耦合的演化规律。

(2) 全球 52 个国家在第一波疫情中的全周期死亡率都符合 Logistic 的演化形式。由此发现了各国死亡率演化早期的非存活者在 ICU 的持续时间 ( $\tau$ ) 与当地新冠疫情的持续时间之间具有普适的线性率关系，二者的比值大约为 7。这是此次新冠疫情展示的一个重要特征，如果正确，人们将能够根据疫情早期的 ICU 临床数据对疫情持续时间做出较为准确的预测。

(3) 死亡率演化晚期的饱和死亡率 ( $m_0$ ) 决定了死亡总人数，而本文发现它与国家中位数年龄、床位占有率为峰值之间分别具有幂律关系。并由此得出结论，若老龄国家 (>38 岁) 的病床数量增加一倍，可减少约 36% 的死亡人数。这一结果定量地解释了各国死亡率具有数量级差异的原因，并且为下一波疫情准备医疗资源提供了定量的依据。

将人体序参量理论应用于人体脉搏波系统，本文主要获得了以下结论：

(1) 本文根据脉搏波基本参数构建了“人体序参量相空间”，以之为基础对亚健康人群划分亚类和计算偏离度，构建了真正符合中医“知常达变”的脉诊原理的定量辨析健康态的模型，突破了传统的大量统计平均学习方法的局限性。

(2) 本文将脉搏波序参量模型应用于人体健康指数的标定，结果显示它能够精确刻画群体稳定的健康指数和个体日变的健康指数，且预测相对误差 (4.1%) 低于传统机器学习的结果 (4.6%)。另外，这套模型框架被应用于刻画其他人体生理波动信号 (如

GDV 信号) 也取得了成功, 显示了方法的可靠性和普适性。

## 7.2 主要创新点

本文主要的创新点是, 在学习继承朗道的“序参量”思想和余振苏等人发展的“结构系综理论”的基础上, 发展了一套基于对称性分析的人体序参量理论, 有望成为人体复杂系统研究的新范式。该范式从对称性分析开始, 通过选择能够刻画人体系统有序度的序参量, 来开展亚类分解和定量建模, 为认识人体复杂系统自组织机制, 突破传统统计平均方法的局限, 开展人体个性化特征的数学建模提供了一种新工具。具体在应用于疫情演化动力学模型和脉搏波分析模型的构建中, 取得了如下进展:

(1) 从序参量理论出发, 本文在疫情早期由病毒主导的 SEIR 传染病动力学模型基础上, 发现并刻画了疫情中晚期由国家干预措施所主导的新的对称破缺过程, 并对此建立了新的 GHR 传染病动力学模型, 从中揭示了新冠疫情传播中普遍存在的包含了“序参量-转换速率-群体数量”三个层次相互耦合的疫情演化规律。

(2) 应用序参量理论, 本文实现了将复杂的新冠死亡率演化行为, 解耦为疫情演化的两个独立阶段。由此, 发现了疫情致死在个体尺度和国家尺度之间的跨尺度关联, 定量地回答了“各国死亡率巨大差异的原因”, 并发现了一种预测“疫情持续时间”的新方法。

(3) 针对人体脉搏波, 本文创建了“人体序参量相空间”, 据此建立了一套完整的划分亚健康人群的方法, 以及计算个体健康指数的模型。这一模型突破了传统大数统计平均方法的局限性, 实现了对中医脉诊个性化诊断原理的定量表达。经过人体脉搏波信号和手指 GDV 信号的检验, 证实了模型的有效性和普适性, 揭示了人体生理波动信号中隐含的高度结构化的特征。这为开展融入医学经验和知识的大规模机器学习提供了理论基础。

## 7.3 研究计划

人体系统的研究充满了复杂性的挑战, 但其支撑的健康产业有着巨大的社会需求。我们将继续在本文研究成果的基础上, 结合健康产业的需求, 针对人体复杂系统开展以下研究:

在新冠疫情演化方面, 针对当前世界疫情的真实发展态势, 有两部分的内容可以开展:

(1) 在 GHR 模型的基础上, 纳入更多类型的人群, 如接种疫苗的人群、无症状感染人群等, 开展更符合当前疫情演化情况的模拟, 总结归纳模型参数的演化规律, 揭示疫情传播和社会干预之间的普适规律。

(2) 在当前死亡率的 Logistic 模型的基础上，模拟多次爆发的疫情国家的死亡率演化行为。并且收集更多传染病或流行病的数据，验证和比较疫情早期非存活者的 ICU 持续时间与疫情持续时间之间的线性规律。由此定量地预测新冠疫情的结束时间。

在人体脉搏波研究方面，针对脉搏波分析的临床应用需求，也有两部分的研究可以开展：

(1) 提炼具体复杂性疾病（如高血压、糖尿病）的脉搏波序参量，构建专门的序参量模型，开展中医证型的划分，定量评估病情的严重程度、康复速度等具有重要临床应用价值的指标。在这个过程中，与临床实践之间多次迭代，完善模型框架。

(2) 研究人体生理波动信号在序参量相空间中的典型结构特征，创新融入医学经验和知识的机器学习。进一步融合其他人体生理波动信号，构建多维信号融合的人体健康态测评系统。



## 参考文献

- [1] 赵光武. 还原论与整体论相结合探索复杂性[J]. 北京大学学报（哲学社会科学版）, 2002, 39(6): 14-19.
- [2] W.海森堡. 物理学与哲学[M]. 范岱年,译. 北京: 商务印书馆, 1981.
- [3] Lorenz E N. Deterministic Nonperiodic Flow [J]. Journal of Atmospheric Sciences, 1963, 20(2): 130-141.
- [4] May R M. Simple mathematical models with very complicated dynamics[J]. Nature, 1976, 261(5560): 459-467.
- [5] 薛定谔. 自然与古希腊[M]. 上海: 上海科学技术出版社, 2002.
- [6] 冯·贝塔朗菲. 一般系统论: 基础、发展和应用[M]. 林康义, 魏宏森等,译. 北京: 清华大学出版社, 1987.
- [7] 苗东升. 复杂性研究的现状与展望[J]. 系统辩证学学报, 2001, 04(04): 3-9.
- [8] 普利高津, 斯唐热. 从混沌到有序:人与自然的新对话[M]. 曾庆宏, 沈小峰,译. 上海: 上海译文出版社, 2005.
- [9] 赫尔曼·哈肯. 协同学: 大自然构成的奥秘[M]. 凌复华,译上海译文出版社, 2005.
- [10] 约翰·H·霍兰. 隐秩序——适应性造就复杂性[M]. 周晓牧, 韩晖,译. 北京: 上海科技教育出版社, 2000.
- [11] 钱学森. 创新系统学 (新世纪版本)[M]. 上海: 上海交通大学出版社, 2007.
- [12] E.拉兹洛, 闵家胤 (译). 从贝塔朗菲的著述看一般系统论的起源 (《关于一般系统论》一书的导言)[J]. 系统辩证学学报, 1993(2): 60-64.
- [13] 陈一壮. 论贝塔朗菲的"一般系统论"与圣菲研究所的"复杂适应系统理论"的区别[J]. 山东科技大学学报: 社会科学版, 2007, 9(2): 5-8.
- [14] 姜璐. “自组织”理论简介[J]. 系统工程, 1986, 4(2): 30-32.
- [15] 沈小峰, 郭治安. 协同学的方法论问题[J]. 北京师范大学学报(自然科学版), 1984, 01: 89-95.
- [16] 胡金星. 产业融合的内在机制研究——基于自组织理论的视角[D]. 复旦大学, 2007.
- [17] 徐全军. 企业理论新探:企业自组织理论[J]. 南开管理评论, 2003, 6(3): 7.
- [18] 李汉卿. 协同治理理论探析[J]. 理论月刊, 2014(1): 5.
- [19] 李伯华, 刘沛林, 窦银娣. 乡村人居环境系统的自组织演化机理研究[J]. 经济地理, 2014, 34(9): 7.
- [20] 姜璐. 自组织理论在社会经济系统中的应用[J]. 系统工程理论与实践, 1988(3): 60-65.
- [21] 苗东升. 开来学于今:复杂性科学纵横论[M]. 开来学于今: 复杂性科学纵横论, 2009.
- [22] Grimm, Volker, Revilla, et al. Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology[J]. Science, 2005, 310(5750): 987-991.
- [23] 余振苏, 倪志勇. 人体复杂系统科学探索[M]. 北京: 科学出版社, 2012.

- [24] Watts D J, Strogatz S H. Collective dynamics of ‘small-world’ networks[J]. *Nature*, 1998, 393(6684): 440-442.
- [25] Barabási A-L, Albert R. Emergence of Scaling in Random Networks[J]. *Science*, 1999, 286(5439): 509-512.
- [26] Motter A E, Albert R. Networks in motion[J]. *Physics Today*, 2012, 65(4): 43-48.
- [27] 祝世讷. 中医是第一门复杂性科学[J]. 山东中医药大学学报, 2016, 40(02): 99-101, 130.
- [28] 祝世讷, 陈少宗. 中医系统论与系统工程学[M]. 北京: 中国医药科技出版社, 2002.
- [29] 中国中医药报社. 哲眼看中医[M]. 北京: 北京科学技术出版社, 2005.
- [30] 于景元. 从系统思想到系统实践的创新——钱学森系统研究的成就和贡献[J]. 系统工程理论与实践, 2016, 36(12): 2993-3002.
- [31] 顾吉珉. 钱学森书信: 一座书信史上的丰碑——从《钱学森书信》和《钱学森书信补编》谈起 [J]. 钱学森研究, 2016, 01: 186-203.
- [32] 余振苏. 复杂系统学新框架-融合量子与道的知识体系[M]. 北京: 科学出版社, 2012.
- [33] 钱学森, 于景元, 戴汝为. 一个科学新领域——开放的复杂巨系统及其方法论[J]. 自然杂志, 1990, 13(01): 3-10+64.
- [34] 于景元. 钱学森综合集成体系[J]. 西安交通大学学报(社会科学版), 2006, 06: 40-47.
- [35] 樊瑛, 狄增如, 方福康. 包含人力资本的宏观经济增长模型[J]. 北京师范大学学报(自然科学版), 2004, 03: 417-421.
- [36] 王众托. 知识系统工程与现代科学技术体系[J]. 上海理工大学学报, 2011, 33(6): 18.
- [37] 操龙兵, 戴汝为. 开放复杂智能系统[M]. 开放复杂智能系统, 2008.
- [38] 顾基发, 王浣尘, 唐锡晋. 综合集成方法体系与系统学研究[M]. 综合集成方法体系与系统学研究, 2007.
- [39] 马蔼乃. 理论地理科学与哲学: 复杂性科学理论[M]. 理论地理科学与哲学: 复杂性科学理论, 2007.
- [40] 马蔼乃. 地理科学导论: 自然科学与社会科学的桥梁科学[M]. 地理科学导论: 自然科学与社会科学的桥梁科学, 2005.
- [41] 马蔼乃. 地理信息科学: 天地人机信息一体化网络系统[M]. 地理信息科学: 天地人机信息一体化网络系统, 2006.
- [42] 马蔼乃. 地理系统工程: 可持续发展战略的基础[M]. 地理系统工程: 可持续发展战略的基础, 2006.
- [43] 马蔼乃. 地理建设与社会系统工程[M]. 地理建设与社会系统工程, 2012.
- [44] 马蔼乃. 地理科学与现代科学技术体系[M]. 地理科学与现代科学技术体系, 2011.
- [45] Barabási, Albert-László. Scale-Free Networks: A Decade and Beyond[J]. *Science*, 2009.
- [46] 郭雷. 系统学是什么[J]. 系统科学与数学, 2016, 3.
- [47] 钱学森, 陈信. 人体科学是现代科学技术体系中的一个大部门[J]. 自然杂志, 1988.
- [48] 钱学森. 论人体科学[M]. 北京: 人民军医出版社, 1988.
- [49] 余振苏, 倪志勇. 论思维的复杂系统模型暨新一代专家系统的设想[J]. 北京大学学报: 自然科学版, 2011, 47(5): 960-968.

- [50] 余振苏, 倪志勇. 钱学森复杂系统思想的理论探索与实践[J]. 党政干部学刊, 2011(1): 8-13.
- [51] 余振苏, 倪志勇, 等. 力学创新助飞奥运梦想——中国激流项目科技攻关纪实[M]. 北京: 科学出版社, 2008.
- [52] Landau L D. Theory of phase transformations. I[J]. Zh. Eksp. Teor. Fiz., 1937, 7: 19.
- [53] Landau L D. Theory of phase transformations. II[J]. Zh. Eksp. Teor. Fiz., 1937, 7: 627.
- [54] 廷哈姆. 超导电性导论[M]. 邵惠民, 宁宇宏, 张其瑞,译. 北京: 科学出版社, 1985.
- [55] The Nobel Prize. Press Release: The Nobel Prize in Physics 2021[EB/OL]. [ 2022-02-24].  
<https://www.nobelprize.org/prizes/physics/2021/press-release/>.
- [56] 余振苏, 程雪玲. 湍流的复杂系统论[C]. 钱学森技术科学思想与力学, 2001.
- [57] She Z-S, Chen X, Hussain F. Quantifying wall turbulence via a symmetry approach: a Lie group theory[J]. Journal of Fluid Mechanics, 2017, 827: 322-356.
- [58] Chen X, Hussain F, She Z-S. Quantifying wall turbulence via a symmetry approach. Part 2. Reynolds stresses[J]. Journal of Fluid Mechanics, 2018, 850: 401-438.
- [59] Li R, She Z-S. Unified energy law for fluctuating density wave orders in cuprate pseudogap phase[J]. Communications Physics, 2022, 5(1): 13.
- [60] 郭蕾, 王永炎, 张俊龙, et al. 阴阳--人体系统序参量解读(一)[J]. 中医药学刊, 2004.
- [61] 郭蕾, 王永炎, 张俊龙, et al. 阴阳——人体系统序参量解读(二)[J]. 中医药学刊, 2004, 22(12): 4.
- [62] 吴邦惠. "气"——人体超复杂巨系统的"序参量"[J]. 中国人体科学, 1992, 2(4): 4.
- [63] 陶必修. 主导人体健康的序参量[J]. 贵州科学, 2015, 33(4): 4.
- [64] 吕嘉戈, 马蔼乃. 气功医学数学模型初探[J]. 中医临床研究, 2018, 10(33): 6.
- [65] Luhao Z, Rong L, Gang S, et al. Impairment of T cells' antiviral and anti-inflammation immunities may be critical to death from COVID-19[J]. Royal Society Open Science, 2021.
- [66] Preparing for the next pandemic[J]. Nature Medicine, 2021, 27(3): 357-357.
- [67] Fredbrauer, Carloscastillo-Chavez, 布劳尔, et al. 生物数学:种群生物学与传染病学中的数学模型 [M]. 生物数学:种群生物学与传染病学中的数学模型, 2013.
- [68] 中华人民共和国国务院新闻办公室. 抗击新冠肺炎疫情的中国行动[EB/OL]. [2022-02-24].  
[http://www.gov.cn/zhengce/2020-06/07/content\\_5517737.htm](http://www.gov.cn/zhengce/2020-06/07/content_5517737.htm).
- [69] 国务院联防联控机制. 国务院联防联控机制权威发布(2020年3月17日) [EB/OL]. [2022-02-24]. <http://www.gov.cn/xinwen/gwylflkjz60/wzsl.htm>.
- [70] Adam D. MODELLING THE PANDEMIC The simulations driving the world's response to COVID-19[J]. Nature, 2020, 580(7803): 316-318.
- [71] Else H. COVID IN PAPERS: A TORRENT OF SCIENCE[J]. Nature, 2020, 588(7839): 553-553.
- [72] Bernoulli D. Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir[J]. Mém Math Phys Acad Roy Sci Paris, 1760: 1-45.
- [73] Kermack W O, Mckendrick A G A. A Contribution to the Mathematical Theory of Epidemics[J]. Proceedings of The Royal Society A Mathematical Physical and Engineering Sciences, 1927, 115(772): 700-721.
- [74] Mahmud A, Lim P Y. Applying the SEIR Model in Forecasting The COVID-19 Trend in Malaysia: A

- Preliminary Study[J]. medRxiv, 2020: 2020.04.14.20065607.
- [75] Labadin J, Hong B H. Transmission Dynamics of COVID-19 in Malaysia Prior to the Movement Control Order[J]. medRxiv, 2020: 2020.02.07.20021188.
- [76] Tang B, Wang X, Li Q, et al. Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions[J]. Journal of Clinical Medicine, 2020, 9(2): 462.
- [77] López L, Rodó X. The end of social confinement and COVID-19 re-emergence risk[J]. Nature Human Behaviour, 2020: 1-10.
- [78] Chen T, Rui J, Wang Q, et al. A mathematical model for simulating the transmission of Wuhan novel Coronavirus[J]. bioRxiv, 2020: 2020.01.19.911669.
- [79] Chen Y, Cheng J, Jiang Y, et al. A time delay dynamical model for outbreak of 2019-nCoV and the parameter identification[J]. Journal of Inverse and Ill-posed Problems, 2020, 28(2): 243-250.
- [80] Yang Z, Zeng Z, Wang K, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions[J]. Journal of Thoracic Disease, 2020.
- [81] Cao S-L, Feng P-H, Shi P-P. Study on the epidemic development of COVID-19 in Hubei province by a modified SEIR model[J]. Journal of Zhejiang University (Medical Science), 2020: 1-13.
- [82] Tianyu Zeng, Yunong Zhang, Zhenyu Li, et al. Predictions of 2019-nCoV Transmission Ending via Comprehensive Methods[J]. arXiv, 2020.
- [83] Biao Tang N L B, Qian Li, Sanyi Tang, Yanni Xiao, Jianhong Wu,. An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov)[J]. Infectious Disease Modelling, 2020, 5: 248-255.
- [84] Peng L, Yang W, Zhang D, et al. Epidemic analysis of COVID-19 in China by dynamical modeling[J]. medRxiv, 2020: 2020.02.16.20023465.
- [85] Godio A, Pace F, Vergnano A. SEIR Modeling of the Italian Epidemic of SARS-CoV-2 Using Computational Swarm Intelligence[J]. International Journal of Environmental Research and Public Health, 2020, 17(10): 3535.
- [86] Zhao Z, Li X, Liu F, et al. Prediction of the COVID-19 spread in African countries and implications for prevention and control: A case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya[J]. Science of The Total Environment, 2020, 729: 138959.
- [87] L. E. Olivier, Craig I K. An epidemiological model for the spread of COVID-19: A South African case study[J]. arXiv, 2020.
- [88] Mangoni L, Pistilli M. Epidemic Analysis of COVID-19 in Italy by Dynamical Modelling[J]. Social Science Electronic Publishing, 2020.
- [89] Cheynet, E. Generalized SEIR Epidemic Model (Fitting and Computation).[EB/OL].  
<https://zenodo.org/record/3911854#collapseCitations>.
- [90] Yang Z, Zeng Z, Wang K, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions[J]. Journal of Thoracic Disease, 2020, 12(3): 165-174.
- [91] Liu X, Huang J, Li C, et al. The role of seasonality in the spread of COVID-19 pandemic - ScienceDirect[J]. Environmental Research, 2021.
- [92] Read J M, Bridgen J R E, Cummings D a T, et al. Novel coronavirus 2019-nCoV (COVID-19): early

- estimation of epidemiological parameters and epidemic size estimates[J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 2021, 376(1829): 20200265.
- [93] Wu J T, Leung K, Leung G M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study[J]. Lancet, 2020, 395(10225): 689-697.
- [94] Ferguson N, Laydon D, Nedjati Gilani G, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand[R]. 2020: 1-20.
- [95] World Health Organization. Tracking SARS-CoV-2 variants[EB/OL]. [2022-03-21]. <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>.
- [96] Yang W, Zhang D, Peng L, et al. Rational evaluation of various epidemic models based on the COVID-19 data of China[J]. Epidemics, 2021, 37.
- [97] Gu Q. Mathematical methods for physics[M]. Beijing: Science Press, 2012.
- [98] 王松桂. 线性统计模型:线性回归与方差分析[M]. 高等教育出版社, 1999.
- [99] Zhang L, She Y-R, She G-H, et al. The cross-scale correlations between individuals and nations in COVID-19 mortality[J]. PREPRINT (Version 2) available at Research Square (under review at Scientific Reports), 2022.
- [100] H Juliette T Unwin, Susan Hillis, Lucie Cluver, et al. Global, regional, and national minimum estimates of children affected by COVID-19-associated orphanhood and caregiver death, by age and family circumstance up to Oct 31, 2021: an updated modelling study[J]. The Lancet Child & Adolescent Health, 2022.
- [101] Mallapaty S. How deadly is the coronavirus? Scientists are close to an answer[J]. Nature, 2020, 582(7813): 467-468.
- [102] Hallal P C, Hartwig F P, Horta B L, et al. Remarkable variability in SARS-CoV-2 antibodies across Brazilian regions: nationwide serological household survey in 27 states[J], 2020.
- [103] Streeck H, Schulte B, Kümmeler B, et al. Infection fatality rate of SARS-CoV-2 in a super-spreading event in Germany[J]. Nature Communications, 2020, 11(1): 5829.
- [104] Viglione G. The true toll of the pandemic[J]. Nature, 2020, 585(7823): 22-24.
- [105] Hartnett K P, Kite-Powell A, Devies J, et al. Impact of the COVID-19 Pandemic on Emergency Department Visits — United States, January 1, 2019–May 30, 2020[J]. MMWR. Morbidity and mortality weekly report, 2020, 69(23).
- [106] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time[J]. Lancet Infectious Diseases, 2020, 20(5): 533-534.
- [107] Wu Z Y, Mcgoogan J M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention[J]. Jama-Journal of the American Medical Association, 2020, 323(13): 1239-1242.
- [108] Chen T, Wu D, Chen H L, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study[J]. Bmj-British Medical Journal, 2020, 368: 12.
- [109] Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel

- Coronavirus-Infected Pneumonia in Wuhan, China[J]. Jama-Journal of the American Medical Association, 2020, 323(11): 1061-1069.
- [110] Perez-Saez J, Lauer S A, Kaiser L, et al. Serology-informed estimates of SARS-CoV-2 infection fatality risk in Geneva, Switzerland[J]. The Lancet. Infectious diseases, 2020.
- [111] Sudharsanan N, Didzun O, Barnighausen T, et al. The Contribution of the Age Distribution of Cases to COVID-19 Case Fatality Across Countries: A 9-Country Demographic Study[J]. Annals of internal medicine, 2020.
- [112] Mallapaty S. The coronavirus is most deadly if you are old and male[J]. Nature, 2020, 585(7823): 16-17.
- [113] Liang L-L, Tseng C-H, Ho H J, et al. Covid-19 mortality is negatively associated with test number and government effectiveness[J]. Scientific Reports, 2020, 10(1).
- [114] Sorci G, Faivre B, Morand S. Explaining among-country variation in COVID-19 case fatality rate[J]. Scientific Reports, 2020, 10(1).
- [115] Khan J R, Awan N, Islam M M, et al. Healthcare Capacity, Health Expenditure, and Civil Society as Predictors of COVID-19 Case Fatalities: A Global Analysis[J]. Frontiers in Public Health, 2020, 8.
- [116] Ji Y, Ma Z, Peppelenbosch M P, et al. Potential association between COVID-19 mortality and healthcare resource availability[J]. Lancet Global Health, 2020, 8(4): E480-E480.
- [117] Kandel N, Chungong S, Omaar A, et al. Health security capacities in the context of COVID-19 outbreak: an analysis of International Health Regulations annual report data from 182 countries[J]. Lancet, 2020, 395(10229): 1047-1053.
- [118] Sen P, Yamana T K, Kandula S, et al. Burden and characteristics of COVID-19 in the United States during 2020[J]. Nature, 2021, 598(7880): 338-341.
- [119] Preventing and controlling SARS-CoV-2 pneumonia outbreaks[EB/OL].  
<https://wjw.hubei.gov.cn/>.
- [120] Roser M, Ritchie H, Ortiz-Ospina E, et al. Coronavirus Pandemic (COVID-19)[J]. Our World in Data, 2020.
- [121] Wuhan Municipal Statistics Bureau, State Statistical Bureau Wuhan Investigation Team. Wuhan Statistical Yearbook 2020[M]. 32. Beijing: China Statistics Press, 2020.
- [122] Sierra N B, Bossuyt N, Braeye T, et al. All-cause mortality supports the COVID-19 mortality in Belgium and comparison with major fatal events of the last century[J]. Archives of Public Health, 2020, 78(1).
- [123] Press Conference[EB/OL]. <http://www.gov.cn/xinwen/gwyflkjz123/index.htm>.
- [124] Herrmann J, Adam E H, Notz Q, et al. COVID-19 Induced Acute Respiratory Distress Syndrome-A Multicenter Observational Study[J]. Frontiers in Medicine, 2020, 7.
- [125] Grasselli G, Greco M, Zanella A, et al. Risk Factors Associated with Mortality among Patients with COVID-19 in Intensive Care Units in Lombardy, Italy[J]. JAMA Internal Medicine, 2020.
- [126] Xie J, Wu W, Li S, et al. Clinical characteristics and outcomes of critically ill patients with novel coronavirus infectious disease (COVID-19) in China: a retrospective multicenter study[J]. Intensive Care Medicine, 2020, 46(10): 1863-1872.

- [127] 杨秦兆虎. 《黄帝内经》述要[M]. 北京: 中国民主法制出版社, 2012.
- [128] 齐向华. 系统辨证脉学培训教程[M]. 北京: 中国中医药出版社, 2022.
- [129] Vierordt K V. Die Lehre vom Arterienpuls in Gesunden und Kranken Zuständen: Gegründet auf eine neue Methode der bildlichen Darstellung des menschlichen Pulses[M]. Braunsweig: Druck Un Verlach Von Friedrich Vieweg and Sohn, 1855.
- [130] Asmar R. 动脉僵硬度和脉搏波速度的临床应用[M]. 王宏宇,译. 北京: 人民军医出版社, 2005.
- [131] .A.麦克唐纳 D. 动脉中的血液流[M]. 柳兆荣,译. 北京: 科学出版社, 1982.
- [132] Williams B, Mancia G, Spiering W, et al. 2018 ESC/ESH Guidelines for the management of arterial hypertension: The Task Force for the management of arterial hypertension of the European Society of Cardiology and the European Society of Hypertension[J]. Journal of Hypertension, 2018, 36.
- [133] Velik R. An objective review of the technological developments for radial pulse diagnosis in Traditional Chinese Medicine[J]. European Journal of Integrative Medicine, 2015, 7(4): 321-331.
- [134] Zhang D-Y. Wrist blood flow signal-based computerized pulse diagnosis using spatial and spectrum features[J]. Journal of Biomedical Science and Engineering, 2010, 03(4): 361-366.
- [135] 费兆馥(主编). 中国脉诊研究[M]. 上海: 上海中医学院出版社, 1991.
- [136] 李永光;张文娟;李德华. 现代脉诊学[M]. 北京: 科学出版社, 2010.
- [137] Wang H, Zhang P. A Quantitative Method for Pulse Strength Classification Based on Decision Tree[C]. International Symposium on Information Science & Engineering, 2008.
- [138] Kim J U, Jeon Y J, Lee Y J, et al. Novel Diagnostic Algorithm for the Floating and Sunken Pulse Qualities and Its Clinical Test[J], 2011, 2011: 813427.
- [139] Shu-Zhong L I, Tian Y, Yang J L, et al. The signal analysis of the slippery pulse and sluggish pulse[J], 2008.
- [140] Xu L, Zhang D, Wang K, et al. Arrhythmic Pulses Detection Using Lempel-Ziv Complexity Analysis[J], 2006, 2006(3): 1-12.
- [141] Zhang D, Zuo W, Zhang D, et al. Classification of Pulse Waveforms Using Edit Distance with Real Penalty[J]. Eurasip Journal on Advances in Signal Processing, 2010.
- [142] Zhao-Xia, Xu, Yi-Qin, et al. Research on Zheng Classification Fusing Pulse Parameters in Coronary Heart Disease[J], 2013, 2013(7): 602672-602679.
- [143] Tang C Y. Developing an objective traditional Chinese medicine pulse diagnostic model in essential hypertension[J], 2010.
- [144] Tang A C Y, Chung J W Y, Wong T K S J E-B C, et al. Validation of a Novel Traditional Chinese Medicine Pulse Diagnostic Model Using an Artificial Neural Network[J], 2011, 2012(1741-427X): 685094.
- [145] Wang X X, Guo P Y J J O B T, University B. TIME-FREQUENCY ANALYSIS OF PULSE SIGNAL BASED ON SYSTEM OF WRIST STRAP SENSOR[J], 2006.
- [146] Xu L, Meng Q H, Qi X, et al. Morphology Variability Analysis of Wrist Pulse Waveform for Assessment of Arteriosclerosis Status[J], 2010, 34(3): 331-339.

- [147] Guo, Wang, Yq, et al. Analysis and Recognition of Traditional Chinese Medicine Pulse Based on the Hilbert-Huang Transform and Random Forest in Patients with Coronary Heart Disease[J], 2015.
- [148] Su Z M, Guo J Y, Liu J J N J Y J G N, et al. Spectrum feature extraction and automatic recognition of pulse waves[J], 2010, 8(1): 70-74.
- [149] Rui G, Yan J, Yan H, et al. TCM Pulse Analysis of the Patient with Coronary Heart Disease Based on Multiscale Entropy[J], 2014, 11th IEEE International Conference on Control & Automation(ICCA).
- [150] Guo R, Wang Y Q, Yan J J, et al. Study of traditional Chinese medicine pulse signals in patients with coronary heart disease based on recurrence quantification analysis[J], 2011, 9(11): 1226-1233.
- [151] Xu W, Yan H, Jin X, et al. Objective study for pulse diagnosis of Traditional Chinese Medicine: Pulse signal analysis of patients with Coronary Heart Disease[C]. 2014 11th IEEE International Conference on Control & Automation (ICCA), 2014.
- [152] Liao Y T, Chen H Y, Huang C M, et al. The pulse spectrum analysis at three stages of pregnancy[J], 2012, 18(4): 382.
- [153] Huang C M, Chang H C, Kao S T, et al. Application of Sphygmography to Detection of Dyspepsia and the Rhinitis[J], 2011, 39(2): 271-285.
- [154] Gong S, Xu B, Sun G, et al. Accurate cirrhosis identification with wrist-pulse data for mobile healthcare[C]. Proceedings of the Second ACM Workshop on Mobile Systems, Applications, and Services for HealthCare, 2012.
- [155] Thakker B, Vyas A L. Frequency domain analysis of radial pulse in abnormal health conditions[C]. Biomedical Engineering & Sciences, 2011.
- [156] Zhang J, Wang R, Lu S, et al. EasiCPRS: Design and implementation of a portable Chinese pulse-wave retrieval system[C]. International Conference on Embedded Networked Sensor Systems, 2011.
- [157] 智能脉诊仪[EB/OL]. [2022-03-27]. <http://www.healthenn.com/yfcg/711.jhtml>.
- [158] 骆璇, 常昌远, 魏同立. 脉搏波的解析模型及初步临床应用研究[J]. 数理医药学杂志, 1999(04): 293-295.
- [159] 常昌远, 骆璇, 缪炳友, et al. 脉搏波的函数叠加逼近分析方法的研究[J]. 应用科学学报, 2000(02): 135-138.
- [160] 钱伟立, 徐兰义, 陈富裕, 郑荣梁. 高斯函数分解法提取脉搏波特征[J]. 中国生物医学工程学报, 1994(01): 1-7+15.
- [161] 周宽久, 陈雪峰. 人体脉搏波特征提取算法研究[C]. 2006 中国控制与决策学术年会, 2006: 5.
- [162] 汪剑鸣, 惠芳, 韦然. 基于特征参数的脉搏波高斯拟合[J]. 天津工业大学学报, 2016(05): 59-65.
- [163] Arthur D, Vassilvitskii S. K-Means++: The Advantages of Careful Seeding[C]. Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, New Orleans, Louisiana, USA, January 7-9, 2007, 2007.
- [164] 刘奕君. “作业和谐”视角下的大学生心理健康问题分析与对策初探[D]. 北京大学, 2021.
- [165] 杨永东, 汤青, 王晶, et al. 基于 GDV 的能量检测装置, 2020.

## 附录 A GDV 图像处理流程

基于新绎健康研究院研发的 GDV 手指辉光采集仪器，可以得到如图 A.1 所示的 GDV 原始图像。

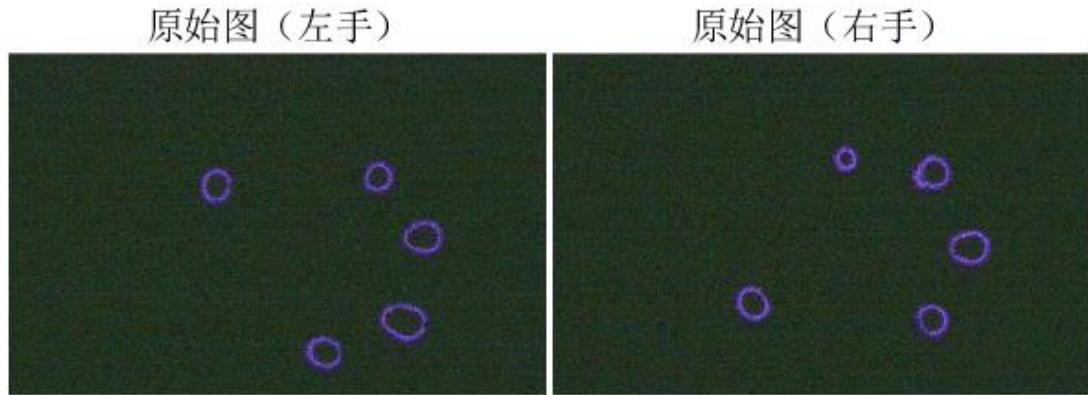


图 A.1 基于新绎 GDV 仪器采集到的 GDV 原始图像。左图为左手图像，右图为右手图像。

本文使用了 MATLAB 软件的 `rgb2gray` 函数将 GDV 的原始 RGB 图像转为了图 A.2 所示的灰度图像，基于此来进行后续处理和计算。这里，`rgb2gray` 函数是通过计算 R、G 和 B 分量的加权和，将 RGB 值转换为灰度值。

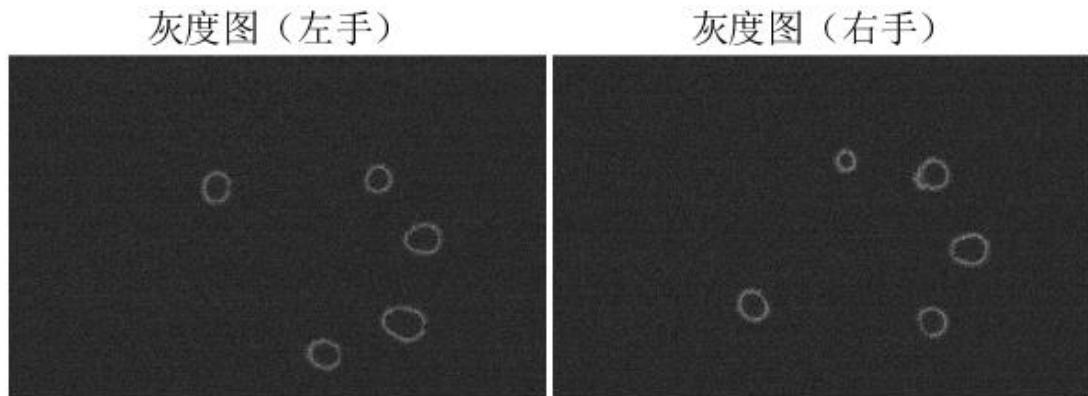


图 A.2 将 GDV 的原始 RGB 图像转为灰度图像。左图为左手图像，右图为右手图像。

进一步，需要将灰度图像转为二值图像，这里，本文使用了 MATLAB 软件的 `imbinarize` 函数将灰度图像转为二值图像。`imbinarize` 函数使用 *Otsu* 方法确定了使得黑白像素的类内方差最小的阈值，通过将所有高于全局阈值的值替换为 1，并将所有其

他值设置为 0，实现了二值图像的转换。为了更好地划分手指区域范围，在转换为二值图之前，需要对灰度图进行滤波降噪预处理。本文使用了二维维纳滤波对图像背景的高斯白噪声进行了滤波处理，滤波前后的二值图结果如图 A.3 所示。

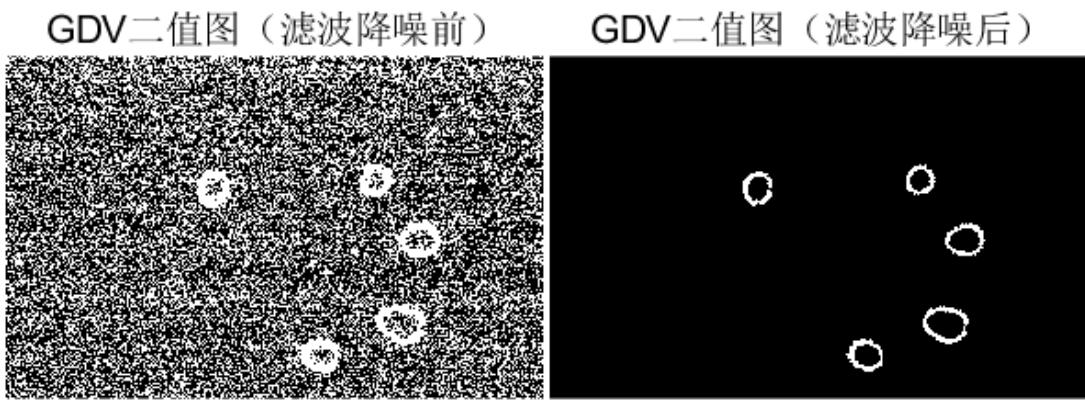


图 A.3 滤波降噪前后的 GDV 二值图像对比。左图为滤波降噪前的图像，右图为滤波降噪处理后的图像。

在上述基础上，需要进一步确定出各手指辉光图像的内外轮廓。由于实际测量的辉光图像存在不同程度的边缘不连续、缺损等情况，本文采用了形态学计算的方法。首先，对 GDV 的二值图进行形态学闭运算，先膨胀图像，然后腐蚀经过膨胀的图像。由此可以填充图像中的小孔，同时保留图像中对象的形状和大小。由此得到 GDV 辉光图像的外轮所包围的面积（见图 A.4b）。进一步根据外轮面积可以提取外轮廓。其次，外轮面积减去辉光图像，即为 GDV 辉光图像的内轮面积（见图 A.4c）。进一步进行形态学开运算，即先腐蚀图像，然后膨胀经过腐蚀的图像，则可以得到较为准确的 GDV 辉光图像的内轮面积（见图 A.4d）。进一步根据内轮面积可以提取内轮廓。

进一步，本文采取了多次聚类迭代的算法，来识别各手指区域，而各手指的聚类中心则为手指的质心。首先，确定的外轮廓面积图，进行 K-means 聚类，初步得到各手指的区域像素点，以及各手指的聚类中心（质心）。

如图 A.5 所示，经过对包含 5 个手指的图像进行全局聚类，得到了初步的各手指区域的像素点以及各手指的质心坐标。但同时也发现，初步聚类的结果有时会把一些背景噪点也划分为手指像素点。例如，在图 A.5 右图顶部的浅蓝色背景噪点，也被归为了拇指区域。如果图像质量较低，背景噪点较多时，这就会对聚类结果和质心坐标产生影响。

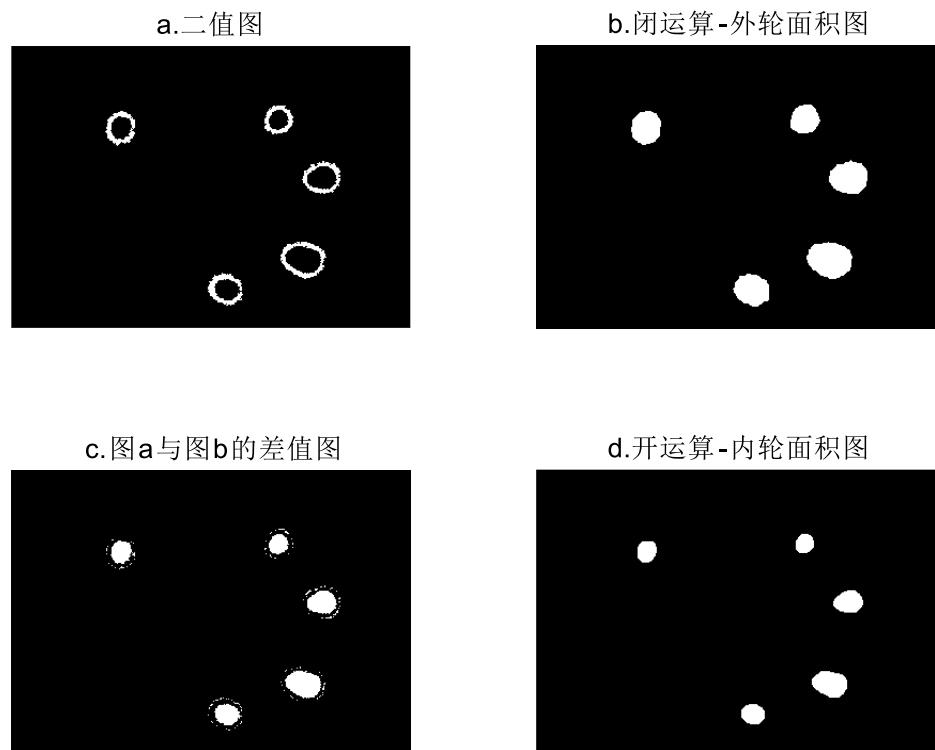


图 A.4 提取辉光内外轮廓的过程。a. 经过滤波降噪处理后的二值图。b. 经过闭运算后的外轮廓面积；c. 外轮廓面积与二值图的差值图。d. 经过开运算后的内轮廓面积图像。

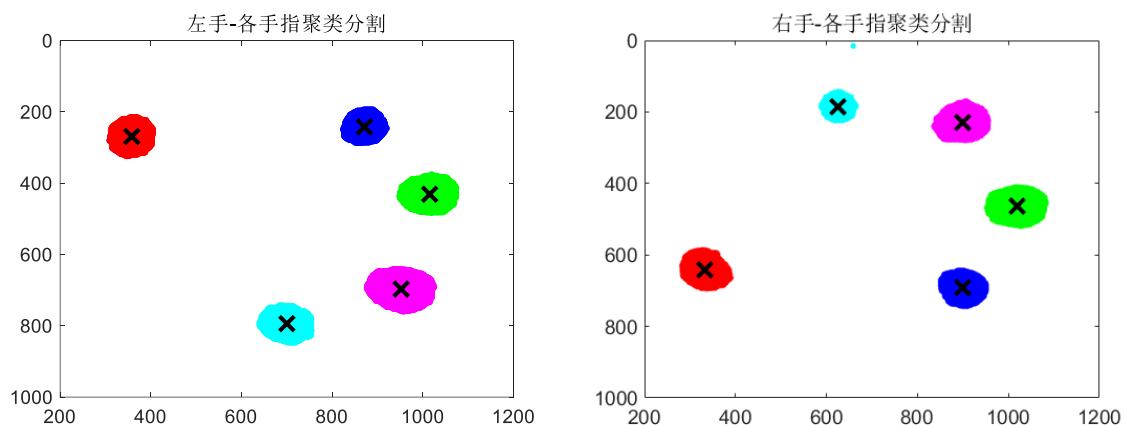


图 A.5 基于聚类算法的 GDV 图像手指识别结果。左图为左手图像，右图为右手图像。红色区域为拇指，深蓝色区域为食指，绿色区域为中指，紫色区域为无名指，浅蓝色区域为小指。

因此，需要在上述各手指区域和质心坐标的基础上，对各手指的内轮廓

分别进行聚类，以获得精确的手指区域和质心坐标。结果如图 A.6 所示。

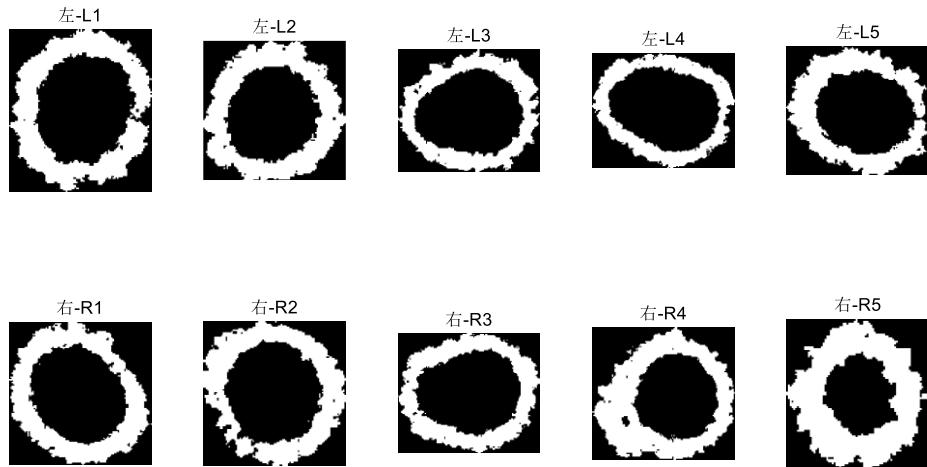


图 A.6 对各手指的内轮廓分别进行聚类后的手指识别结果。上图为左手图像，下图为右手图像。1-5 分别代表拇指、食指、中指、无名指和小指。

进一步，就可以开展对各手指辉光的亮度和形状等参数进行提取和计算。根据确定出的各手指的区域范围，我们可以划分出其所对应的二值图像和灰度图像（未经过滤波降噪）。如图 A.7 所示。

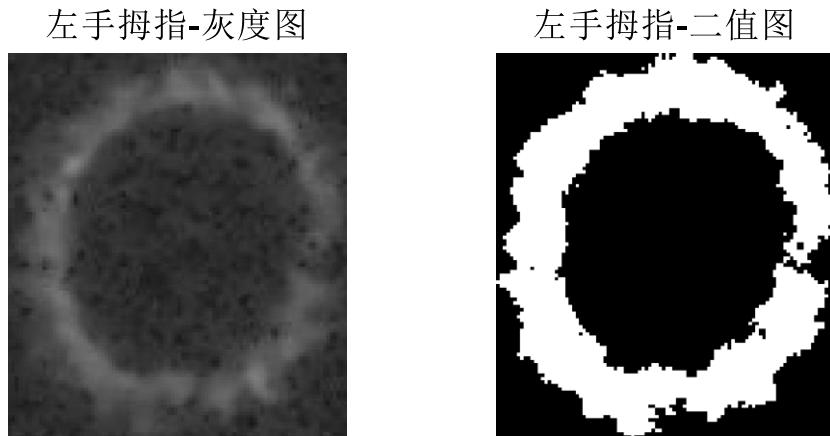


图 A.7 左手拇指的灰度图和二值图。左图为灰度图，右图为二值图。

由此，可以分别得到灰度图直方图，以及二值图所确定的辉光的灰度图直方图，如图 A.8 所示。灰度直方图主要分为两部分，一部分是以较暗（灰度值较低）的背景噪声为主的蓝色区域，一部分是以较亮（灰度值较高）的辉光为主的黄色区域。通过高斯函

数拟合蓝色区域主波的上升沿和下降沿，则可以获得背景高斯白噪声所代表的区域（红色曲线所围区域）。进一步，蓝色区域减去红色区域后，剩下的区域则为辉光所对应的灰度范围（黑色曲线所围区域）。这种方式获得的辉光范围，有效地避免了使用全局阈值划分区域所带来的辉光边缘较暗部分的信息丢失。

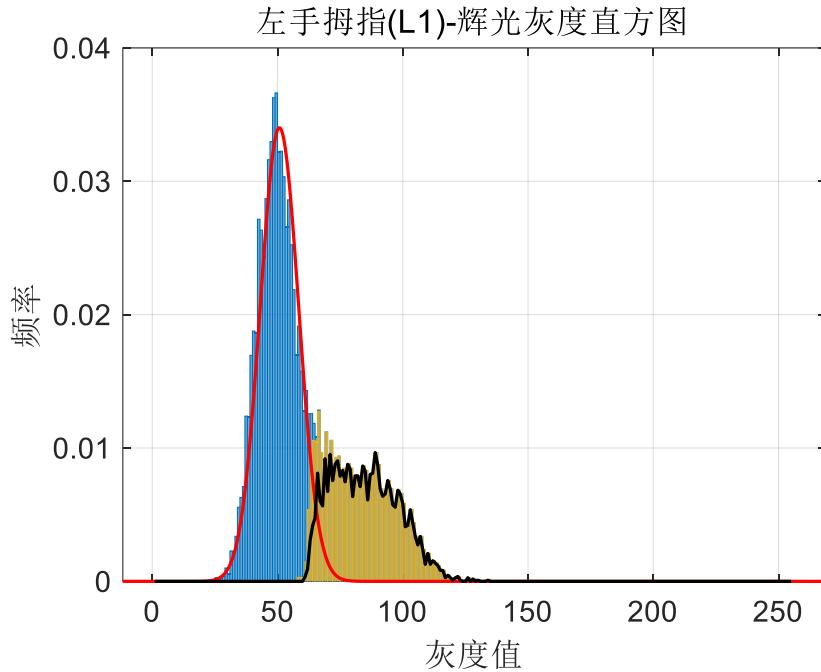


图 A.8 左手拇指的灰度直方图。图中横坐标为图像的灰度值（0-255），纵坐标是对应灰度出现的频率。蓝色区域为整个灰度图的频率直方图，黄色区域是辉光范围对应的频率直方图。红色曲线是用高斯函数拟合蓝色区域的主波上升沿和下降沿后得到的拟合曲线，黑色曲线围成的区域是蓝色区域减去红色曲线包围面积之后剩下的区域。

最终我们提取了两个量，十指平均亮度 $\bar{I}$ :

$$\bar{I} = \frac{1}{10} \sum_{j=1}^{10} \sum_{i=1}^{I_{range}} I_i * f_i \quad (\text{A.1})$$

以及十指平均面积 $\bar{S}$ :

$$\bar{S} = \frac{1}{10} \sum_{j=1}^{10} \sum_{i=1}^{I_{range}} f_i \quad (\text{A.2})$$

其中， $I_{range} = I_{max} - I_{min}$ ，表示 GDV 信号灰度直方图上黑色曲线的 x 轴的坐标范围。



## 附录 B 脉搏波与健康指数的原始数据

表 B.1 六部脉位 (P) 的原始数据

日期	脉位-基线压力 P (g)					
	左寸	左关	左尺	右寸	右关	右尺
2019-11-03	167.35	136.74	182.27	250.92	183.46	243.81
2019-11-04	165.69	189.33	238.72	236.59	168.09	247.51
2019-11-05	218.18	139.85	209.45	168.97	198.50	254.15
2019-11-06	156.18	161.87	211.23	250.27	192.33	245.24
2019-11-08	246.33	139.25	217.27	243.83	162.06	213.58
2019-11-10	246.11	116.28	215.69	249.43	163.10	231.95
2019-11-12	219.24	213.46	241.63	246.91	161.26	189.88
2019-11-13	242.81	117.00	190.15	253.91	138.73	196.21
2019-11-17	135.03	217.17	229.97	232.98	182.24	188.06
2019-11-18	121.33	186.03	232.85	217.02	142.83	212.61
2019-11-19	115.17	161.52	233.67	197.47	111.92	166.94
2019-11-20	116.41	142.33	231.33	191.17	122.51	175.75
2019-11-21	112.31	134.50	210.46	140.29	107.21	138.97
2019-11-26	245.44	110.99	185.00	244.59	134.51	144.38
2019-11-27	252.04	187.12	193.24	252.34	142.94	249.40
2019-11-28	247.45	193.20	190.46	249.36	138.69	245.34
2019-11-29	244.38	145.35	238.91	246.52	223.32	247.69
2019-12-01	186.08	217.58	241.51	249.39	218.14	196.18
2019-12-02	195.41	112.13	193.19	249.52	148.92	208.12
2019-12-03	243.34	217.83	235.15	238.39	245.76	233.98
2019-12-04	250.61	114.51	191.98	243.86	162.82	215.87
2019-12-05	244.31	121.55	182.60	246.17	245.32	211.47
2019-12-06	243.81	123.05	167.64	253.77	221.12	214.49

## 北京大学博士学位论文

日期	脉位-基线压力 P (g)					
	左寸	左关	左尺	右寸	右关	右尺
2019-12-07	238.13	151.73	191.23	255.75	234.63	246.11
2019-12-09	187.64	163.31	214.08	246.36	194.78	143.78
2019-12-11	249.87	142.43	188.93	224.63	199.98	194.31
2019-12-12	239.72	137.76	181.75	223.83	141.85	212.57
2019-12-13	251.30	143.46	187.23	261.82	141.25	184.85
2019-12-14	236.11	164.05	182.61	232.90	114.82	143.07
2019-12-15	235.53	173.04	187.74	250.14	159.03	238.98
2019-12-16	148.27	116.88	161.41	211.60	178.72	242.85
2019-12-17	157.11	115.18	166.45	239.80	163.21	248.16
2019-12-18	243.40	114.51	185.18	240.15	168.23	248.54
2019-12-19	239.05	117.41	185.04	256.83	209.35	175.64
2019-12-20	192.71	114.31	215.54	247.36	176.54	212.73
2019-12-22	238.89	95.86	198.39	225.01	194.60	197.77
2019-12-23	142.55	138.21	192.25	234.66	182.82	243.19
2019-12-24	167.64	135.74	219.03	247.83	139.63	232.03
2019-12-25	246.95	117.29	182.07	248.82	148.61	203.19
2019-12-26	165.13	135.48	166.05	228.49	185.08	241.45
2019-12-30	130.68	118.59	182.20	249.80	143.12	229.56
2019-12-31	234.78	193.88	190.58	241.72	134.91	219.97
2020-01-01	190.45	137.59	160.04	246.76	122.18	185.73
2020-01-02	220.72	88.83	231.45	244.98	147.53	146.85
2020-01-03	187.10	120.76	184.13	244.95	143.20	242.49
2020-01-04	235.93	131.01	205.85	242.41	134.92	165.80
2020-01-06	244.16	145.25	248.90	251.14	164.18	238.19
2020-01-07	212.61	117.78	236.18	243.58	184.60	242.69
2020-01-08	213.32	142.27	191.06	247.08	120.16	245.60
2020-01-09	244.73	138.85	237.26	242.90	218.55	228.51
2020-01-10	247.35	142.25	208.81	226.89	169.10	240.17

附录 B 脉搏波与健康指数的原始数据

日期	脉位-基线压力 P (g)					
	左寸	左关	左尺	右寸	右关	右尺
2020-01-12	233.92	189.47	161.61	220.01	237.13	235.47
2020-01-13	211.72	163.34	192.85	251.36	246.93	210.88
2020-01-15	207.99	173.97	198.54	222.80	187.60	232.79
2020-01-16	213.23	167.97	167.96	242.38	203.17	168.63
2020-01-17	250.23	113.80	184.83	169.27	238.54	209.71
2020-01-19	113.61	147.99	219.05	241.63	190.08	193.38
2020-01-20	172.23	136.08	183.28	238.92	160.92	242.24
2020-01-21	170.69	170.14	237.35	252.49	182.00	239.18
2020-01-22	136.95	138.30	182.29	174.12	246.10	195.60
2020-01-23	215.93	200.62	183.91	242.53	250.01	241.97
2020-01-26	214.69	128.55	239.73	242.51	240.85	243.79
2020-01-27	231.40	190.95	240.18	240.72	245.52	244.64
2020-01-28	242.26	230.75	239.76	220.51	218.05	244.53
2020-01-29	205.42	221.78	237.87	255.59	245.26	236.75
2020-01-30	250.07	212.15	245.97	248.76	244.69	221.04
2020-01-31	241.61	212.79	237.32	247.23	254.06	239.36
2020-02-03	247.61	193.02	243.04	254.79	247.61	246.45
2020-02-04	223.47	149.70	239.25	240.88	215.99	240.21
2020-02-05	236.53	217.09	235.10	249.04	231.03	248.54
2020-02-06	228.58	181.48	237.72	239.09	175.34	247.83
2020-02-07	247.57	240.99	242.28	251.53	239.53	237.52
2020-02-09	252.91	214.05	233.33	241.19	187.49	230.64
2020-02-10	244.14	188.79	245.34	248.95	232.50	242.57
2020-02-11	247.52	211.62	241.27	245.60	194.59	243.45
2020-02-12	242.38	193.40	239.96	246.62	199.86	245.65
2020-02-14	247.38	141.15	215.41	246.11	219.69	232.73
2020-02-15	245.24	145.10	223.75	225.64	222.26	218.46
2020-02-16	241.02	196.98	211.74	248.89	217.85	212.83

日期	脉位-基线压力 P (g)					
	左寸	左关	左尺	右寸	右关	右尺
2020-02-17	141.07	212.03	241.10	241.08	241.88	242.49
2020-02-18	217.67	141.18	188.74	212.15	191.10	238.66
2020-02-19	242.35	171.63	241.66	167.60	224.20	212.04
2020-02-20	166.05	188.88	239.34	243.68	202.65	235.59
2020-02-21	193.55	134.06	239.86	246.38	242.89	235.30
2020-02-22	155.23	145.64	230.85	261.73	216.78	242.45
2020-02-23	189.46	139.17	238.11	240.68	208.72	242.51
2020-02-25	135.86	193.54	242.60	243.14	237.48	242.20
2020-02-26	193.85	134.04	211.91	253.78	195.72	212.75
2020-02-27	174.97	141.08	238.08	171.93	193.45	214.55
2020-02-28	218.32	246.10	244.17	222.50	251.71	241.49

表 B.2 六部脉力 (F) 的原始数据

日期	脉力-主波振幅 F (g)					
	左寸	左关	左尺	右寸	右关	右尺
2019-11-03	29.93	29.87	36.07	6.38	23.84	26.69
2019-11-04	29.60	24.89	36.50	21.81	24.38	13.65
2019-11-05	25.98	26.96	33.24	24.78	26.38	33.51
2019-11-06	31.37	24.47	41.80	18.44	27.17	32.49
2019-11-08	25.88	26.91	31.36	14.92	24.64	33.55
2019-11-10	34.08	23.68	35.36	19.51	27.40	32.71
2019-11-12	23.14	34.95	30.42	29.91	28.39	29.29
2019-11-13	38.18	26.24	32.03	17.13	28.26	29.26
2019-11-17	33.99	35.35	40.78	33.49	27.55	29.83
2019-11-18	32.32	31.49	39.03	30.29	27.80	28.44
2019-11-19	26.37	27.83	36.11	34.53	25.74	26.59
2019-11-20	31.98	26.71	37.73	39.13	27.10	26.51
2019-11-21	33.14	26.10	35.65	33.72	28.40	24.91

附录 B 脉搏波与健康指数的原始数据

日期	脉力-主波振幅 F (g)					
	左寸	左关	左尺	右寸	右关	右尺
2019-11-26	28.90	24.54	31.51	28.06	21.49	17.94
2019-11-27	9.43	25.20	22.41	6.79	19.11	20.46
2019-11-28	19.93	17.46	23.48	15.78	20.22	23.06
2019-11-29	10.25	16.62	22.17	8.14	15.57	20.52
2019-12-01	24.91	17.98	26.18	11.13	15.69	17.61
2019-12-02	25.42	21.94	25.49	18.68	18.44	27.94
2019-12-03	30.08	24.88	29.10	19.59	23.54	23.51
2019-12-04	24.80	22.97	33.56	16.97	27.24	27.14
2019-12-05	29.90	20.37	30.09	12.06	20.69	24.43
2019-12-06	23.48	19.14	25.19	8.57	23.58	22.78
2019-12-07	31.54	29.88	26.04	7.82	25.99	24.53
2019-12-09	34.57	27.52	37.20	22.95	27.83	26.71
2019-12-11	20.77	20.15	30.38	24.85	23.83	26.37
2019-12-12	23.77	29.08	36.64	27.63	26.30	31.07
2019-12-13	28.40	22.49	29.60	21.74	23.22	30.74
2019-12-14	30.15	26.03	32.72	24.44	27.08	23.21
2019-12-15	28.99	26.75	32.89	15.34	26.92	16.54
2019-12-16	31.20	24.69	28.40	16.26	26.15	27.79
2019-12-17	29.21	24.91	28.42	22.92	27.32	25.15
2019-12-18	33.14	24.08	37.24	18.11	25.99	17.11
2019-12-19	34.92	27.52	29.59	25.31	27.91	28.35
2019-12-20	25.57	20.78	31.48	13.16	24.23	30.26
2019-12-22	23.84	21.01	27.91	14.88	28.24	27.67
2019-12-23	21.27	20.64	32.58	23.58	25.20	32.10
2019-12-24	27.51	27.70	32.96	9.43	22.09	28.91
2019-12-25	23.90	18.87	30.35	10.27	27.16	31.43
2019-12-26	32.78	29.18	31.57	8.62	26.73	21.53
2019-12-30	29.54	30.77	33.21	16.65	29.35	30.05

## 北京大学博士学位论文

日期	脉力-主波振幅 F (g)					
	左寸	左关	左尺	右寸	右关	右尺
2019-12-31	25.51	24.35	35.28	16.76	27.50	33.31
2020-01-01	36.16	25.21	25.39	27.18	21.25	29.93
2020-01-02	26.49	26.98	37.42	12.49	28.74	29.58
2020-01-03	32.29	25.14	32.90	15.51	28.71	30.96
2020-01-04	36.72	28.82	35.79	16.36	26.64	30.31
2020-01-06	29.95	28.78	31.98	23.07	24.77	27.02
2020-01-07	34.01	21.88	34.19	23.36	28.25	27.19
2020-01-08	28.55	27.32	30.97	12.27	26.72	25.04
2020-01-09	28.13	23.81	31.13	16.54	26.84	33.59
2020-01-10	28.22	23.88	29.25	7.77	26.40	28.49
2020-01-12	27.71	28.53	23.25	11.51	19.48	22.74
2020-01-13	41.03	33.97	37.47	21.21	18.19	23.10
2020-01-15	34.62	20.90	33.02	26.15	24.70	24.41
2020-01-16	34.96	27.72	28.46	23.03	27.25	29.13
2020-01-17	27.76	23.84	28.36	22.13	27.96	30.27
2020-01-19	26.50	27.41	37.87	25.47	29.33	29.71
2020-01-20	31.57	23.11	28.65	16.29	29.43	33.87
2020-01-21	22.89	25.46	29.59	14.14	28.11	35.19
2020-01-22	33.79	28.32	37.74	7.88	24.64	32.61
2020-01-23	35.57	33.38	36.28	22.66	28.79	23.59
2020-01-26	30.48	4.39	32.69	16.13	16.68	26.16
2020-01-27	26.33	19.31	22.40	13.32	15.35	21.54
2020-01-28	6.77	16.70	29.34	16.49	26.70	14.63
2020-01-29	35.02	16.41	28.06	10.77	23.87	24.56
2020-01-30	20.70	18.71	32.67	5.82	10.45	11.69
2020-01-31	20.52	15.63	23.89	19.10	19.38	16.64
2020-02-03	17.15	21.20	8.00	6.25	6.20	16.43
2020-02-04	29.25	22.23	21.71	27.09	24.30	29.54

## 附录 B 脉搏波与健康指数的原始数据

日期	脉力-主波振幅 F (g)					
	左寸	左关	左尺	右寸	右关	右尺
2020-02-05	25.37	13.18	36.32	20.09	11.19	17.77
2020-02-06	23.85	19.38	15.78	19.79	20.90	8.33
2020-02-07	13.83	13.21	25.18	16.75	20.09	26.33
2020-02-09	24.86	18.02	34.85	16.19	25.65	22.44
2020-02-10	22.70	22.84	31.42	12.21	21.65	15.95
2020-02-11	21.26	16.78	20.82	8.55	6.28	9.62
2020-02-12	17.72	22.93	27.35	9.56	20.83	14.21
2020-02-14	23.92	23.19	31.89	15.45	23.22	16.53
2020-02-15	17.11	21.33	32.31	10.36	19.13	27.87
2020-02-16	16.10	18.59	34.40	10.34	23.20	18.82
2020-02-17	24.90	22.82	37.49	17.59	9.03	12.46
2020-02-18	16.18	22.33	29.06	27.23	25.35	28.74
2020-02-19	14.81	21.86	20.44	20.41	20.89	24.55
2020-02-20	23.22	21.64	31.22	18.55	20.42	23.23
2020-02-21	18.58	22.22	24.87	6.99	19.45	25.40
2020-02-22	27.27	21.47	33.71	9.67	22.07	21.41
2020-02-23	24.89	24.92	31.34	22.04	21.18	23.26
2020-02-25	29.69	29.82	26.25	22.27	20.70	29.39
2020-02-26	25.34	23.61	34.43	6.94	20.17	27.76
2020-02-27	32.51	27.39	31.84	26.22	25.23	30.00
2020-02-28	23.37	18.47	35.98	7.22	19.10	15.43

表 B.3 六部周期 (T) 的原始数据

日期	周期-搏动时间 T (s)					
	左寸	左关	左尺	右寸	右关	右尺
2019-11-03	0.83	0.94	0.95	0.93	1.00	0.91
2019-11-04	0.97	0.93	0.87	0.92	0.93	0.91
2019-11-05	0.92	0.92	0.91	0.94	0.94	0.91

日期	周期-搏动时间 T (s)					
	左寸	左关	左尺	右寸	右关	右尺
2019-11-06	0.98	0.91	0.92	0.93	0.93	0.87
2019-11-08	0.75	0.73	0.62	0.68	0.69	0.69
2019-11-10	0.78	0.74	0.77	0.77	0.83	0.75
2019-11-12	0.85	0.82	0.78	0.86	0.87	0.85
2019-11-13	0.80	0.81	0.84	0.81	0.79	0.84
2019-11-17	0.85	0.85	0.82	0.74	0.80	0.78
2019-11-18	0.81	0.79	0.82	0.75	0.76	0.76
2019-11-19	0.78	0.76	0.76	0.75	0.74	0.80
2019-11-20	0.68	0.63	0.65	0.60	0.64	0.62
2019-11-21	0.66	0.67	0.68	0.63	0.66	0.68
2019-11-26	0.68	0.73	0.79	0.73	0.74	0.75
2019-11-27	0.56	0.54	0.57	0.57	0.59	0.52
2019-11-28	0.50	0.49	0.52	0.53	0.51	0.51
2019-11-29	0.59	0.64	0.71	0.57	0.61	0.60
2019-12-01	0.81	0.79	0.77	0.69	0.67	0.78
2019-12-02	0.74	0.74	0.71	0.75	0.80	0.77
2019-12-03	0.70	0.72	0.68	0.71	0.73	0.72
2019-12-04	0.83	0.77	0.83	0.75	0.78	0.72
2019-12-05	0.66	0.66	0.69	0.68	0.67	0.64
2019-12-06	0.73	0.79	0.80	0.74	0.77	0.78
2019-12-07	0.80	0.78	0.79	0.67	0.66	0.67
2019-12-09	0.80	0.78	0.74	0.79	0.73	0.79
2019-12-11	0.70	0.77	0.78	0.80	0.81	0.83
2019-12-12	0.76	0.84	0.81	0.85	0.82	0.81
2019-12-13	0.79	0.82	0.79	0.79	0.82	0.81
2019-12-14	0.71	0.74	0.81	0.78	0.84	0.83
2019-12-15	0.63	0.76	0.76	0.71	0.76	0.81
2019-12-16	0.79	0.79	0.79	0.81	0.72	0.74

附录 B 脉搏波与健康指数的原始数据

日期	周期-搏动时间 T (s)					
	左寸	左关	左尺	右寸	右关	右尺
2019-12-17	0.70	0.74	0.72	0.70	0.75	0.77
2019-12-18	0.82	0.80	0.79	0.77	0.74	0.72
2019-12-19	0.93	0.91	0.89	0.89	0.84	0.80
2019-12-20	1.00	1.01	1.01	0.96	1.02	0.88
2019-12-22	0.66	0.75	0.74	0.76	0.75	0.72
2019-12-23	0.70	0.97	0.87	0.91	0.89	0.95
2019-12-24	0.88	0.88	0.87	0.79	0.86	0.89
2019-12-25	0.89	0.87	0.79	0.84	0.88	0.90
2019-12-26	0.72	0.70	0.82	0.78	0.87	0.87
2019-12-30	0.91	0.91	0.96	0.91	0.91	0.91
2019-12-31	0.92	0.96	0.96	0.92	0.89	0.97
2020-01-01	0.93	0.94	0.90	0.94	0.90	0.94
2020-01-02	0.82	0.73	0.84	0.76	0.77	0.77
2020-01-03	0.97	1.06	1.05	1.07	1.02	0.98
2020-01-04	0.93	0.86	0.83	0.84	0.83	0.82
2020-01-06	0.91	0.93	0.92	0.89	0.97	0.97
2020-01-07	0.69	0.56	0.75	0.78	0.74	0.71
2020-01-08	0.82	0.79	0.75	0.84	0.73	0.76
2020-01-09	0.96	0.93	0.89	0.84	0.93	1.02
2020-01-10	0.89	0.88	0.90	0.88	0.96	0.97
2020-01-12	0.92	0.89	0.89	0.90	0.92	0.91
2020-01-13	0.86	0.85	0.90	0.89	0.93	0.89
2020-01-15	0.93	0.86	0.91	0.84	0.90	0.84
2020-01-16	0.79	0.73	0.81	0.75	0.77	0.81
2020-01-17	0.85	0.85	0.83	0.82	0.80	0.88
2020-01-19	0.76	0.71	0.70	0.75	0.77	0.75
2020-01-20	0.75	0.74	0.71	0.75	0.75	0.70
2020-01-21	0.89	0.83	0.84	0.81	0.83	0.80

日期	周期-搏动时间 T (s)					
	左寸	左关	左尺	右寸	右关	右尺
2020-01-22	0.80	0.76	0.80	0.83	0.81	0.78
2020-01-23	0.66	0.62	0.68	0.66	0.67	0.61
2020-01-26	0.86	0.85	0.82	0.83	0.74	0.82
2020-01-27	0.85	0.89	0.80	0.73	0.74	0.78
2020-01-28	0.80	0.82	0.83	0.81	0.82	0.77
2020-01-29	0.85	0.86	0.82	0.87	0.84	0.85
2020-01-30	0.83	0.88	0.87	0.83	0.85	0.90
2020-01-31	0.80	0.80	0.71	0.77	0.77	0.67
2020-02-03	0.88	0.84	0.82	0.80	0.71	0.74
2020-02-04	0.72	0.83	0.74	0.80	0.82	0.79
2020-02-05	0.98	0.93	0.92	0.86	0.87	0.86
2020-02-06	0.94	0.94	0.94	0.90	0.95	0.97
2020-02-07	0.85	0.90	0.93	0.79	0.78	0.81
2020-02-09	0.79	0.84	0.88	0.82	0.86	0.80
2020-02-10	0.93	0.97	0.92	0.91	0.90	0.91
2020-02-11	0.91	0.88	0.95	0.93	0.88	0.88
2020-02-12	0.98	0.94	0.87	0.93	0.92	0.88
2020-02-14	0.89	0.86	0.90	0.88	0.90	0.85
2020-02-15	0.91	0.90	0.91	0.87	0.96	0.94
2020-02-16	0.77	0.75	0.79	0.81	0.85	0.83
2020-02-17	0.95	0.91	0.89	0.93	0.93	0.91
2020-02-18	0.95	0.92	0.92	0.89	0.91	0.84
2020-02-19	0.95	0.97	0.89	0.92	0.89	0.83
2020-02-20	0.88	0.94	0.87	0.84	0.90	0.88
2020-02-21	0.96	1.00	0.95	0.92	0.83	0.87
2020-02-22	0.88	0.90	0.88	0.92	0.86	0.92
2020-02-23	0.86	0.96	0.94	0.92	0.91	0.91
2020-02-25	0.91	0.93	0.89	0.86	0.92	0.83

## 附录 B 脉搏波与健康指数的原始数据

日期	周期-搏动时间 T (s)					
	左寸	左关	左尺	右寸	右关	右尺
2020-02-26	0.93	0.91	0.89	0.92	0.85	0.86
2020-02-27	0.90	0.89	0.88	0.96	0.90	0.85
2020-02-28	0.97	1.02	0.91	0.93	0.93	0.91

表 B.4 健康指数 (H) 的原始数据

日期	健康指数	日期	健康指数	日期	健康指数
2019-11-03	91.23	2019-12-16	95.61	2020-01-23	95.61
2019-11-04	93.86	2019-12-17	95.61	2020-01-26	94.74
2019-11-05	88.60	2019-12-18	92.11	2020-01-27	95.61
2019-11-06	92.11	2019-12-19	94.74	2020-01-28	94.74
2019-11-08	92.98	2019-12-20	94.74	2020-01-29	96.49
2019-11-10	93.86	2019-12-22	94.74	2020-01-30	94.74
2019-11-12	94.74	2019-12-23	95.61	2020-01-31	94.74
2019-11-13	91.23	2019-12-24	96.49	2020-02-03	96.49
2019-11-17	98.25	2019-12-25	96.49	2020-02-04	96.49
2019-11-18	94.74	2019-12-26	95.61	2020-02-05	96.49
2019-11-19	94.74	2019-12-30	94.74	2020-02-06	93.86
2019-11-20	92.11	2019-12-31	95.61	2020-02-07	97.37
2019-11-21	92.98	2020-01-01	94.74	2020-02-09	96.49
2019-11-26	94.74	2020-01-02	92.98	2020-02-10	96.49
2019-11-27	90.35	2020-01-03	91.23	2020-02-11	94.74
2019-11-28	91.23	2020-01-04	94.74	2020-02-12	96.49
2019-11-29	91.23	2020-01-06	93.86	2020-02-14	97.37
2019-12-01	92.11	2020-01-07	92.98	2020-02-15	95.61
2019-12-02	92.98	2020-01-08	93.86	2020-02-16	94.74
2019-12-03	96.49	2020-01-09	95.61	2020-02-17	98.25
2019-12-04	94.74	2020-01-10	94.74	2020-02-18	92.11
2019-12-05	98.25	2020-01-12	94.74	2020-02-19	97.37
2019-12-06	94.74	2020-01-13	94.74	2020-02-20	97.37
2019-12-07	97.37	2020-01-15	94.74	2020-02-21	98.25
2019-12-09	95.61	2020-01-16	94.74	2020-02-22	93.86

北京大学博士学位论文

2019-12-11	96.49	2020-01-17	93.86	2020-02-23	95.61
2019-12-12	96.49	2020-01-19	93.86	2020-02-25	97.37
2019-12-13	96.49	2020-01-20	94.74	2020-02-26	96.49
2019-12-14	94.74	2020-01-21	95.61	2020-02-27	97.37
2019-12-15	97.37	2020-01-22	95.61	2020-02-28	100.00

## 个人简历、在学期间发表的学术论文与研究成果

### 个人简历

1991年9月出生于甘肃省兰州市。

2008年9月考入中国地质大学（北京）工程技术学院勘查技术与工程专业，2012年7月本科毕业并获得工学学士学位。

2012年9月保送中国地质大学（北京）工程技术学院地质工程专业，2015年7月硕士毕业并获得工学硕士学位，师从孙春岩教授。

2015年9月至2018年8月，入职北京大学工学院健康系统工程研究所，担任科研助理。

2018年9月考入北京大学攻读力学（生物力学与医学工程）博士学位，师从余振苏教授。

### 科研成果

#### 论文

- [1] **Zhang L**, She Y-R, She G-H, Li R, She Z-S. The cross-scale correlations between individuals and nations in COVID-19 mortality[J]. Under review at Scientific Reports, 2021.
- [2] **Zhang L**, She G-H, She Y-R, Li R, She Z-S. A novel symmetry-based dynamics model to quantify the state interventions for COVID-19[J]. in preparation, 2022.
- [3] **Zhang L<sup>\*</sup>**, She Y-R, Xin Y, Zhao Q-C. A TCM principle-based model of order parameters for pulse wave diagnosis[J]. in preparation, 2022.
- [4] **Zhang L**, She Y-R, Xin Y, She Z-S. A TCM principle-based model of pulse wave order parameters to quantify daily fluctuation of individual health status[J]. in preparation, 2022.
- [5] **Zhang L**, Xin Y, Zhao Q-C, She Z-S. A TCM principle-based model of order parameters for gas discharge visualization[J]. in preparation, 2022.

#### 专利

- [1] 宋舒娜, 张磊, 余钰蝶, 俞燕明. 一种基于人体序参量模型的脉搏波智能辨析方法及系统[P](申请号: PCT/CN2022/094000), 2022-05-22.

- [2] 宋舒娜, 张磊,余钰嵘. 一种基于脉搏波圆运动模型的体质辨析方法及系统[P](申请号: 202210557221.3), 2022-05-22.
- [3] 宋舒娜, 张磊, 余钰嵘, 赵前程, 余睿麒, 余国婷. 一种基于脉搏波序参量相空间的情志辨析方法及系统[P](申请号: 202210557336.2), 2022-05-22.
- [4] 宋舒娜, 余睿麒, 宋景和, 张磊, 余光会. 一种个性化作业时长智能分析方法及系统[P](申请号: 202210557340.9), 2022-05-22.
- [5] 宋舒娜, 余睿麒, 赵前程, 余钰嵘, 宋景和, 张磊, 余光会. 一种基于五维模型的身心健康态智能测评方法及系统[P](申请号: 202210558112.3), 2022-05-22.

## 会议

- [1] 张磊, 余振苏. 人体健康的复杂系统论建构(墙报) [C], 第一届中国系统科学大会, 北京, 2017.5.13.
- [2] 张磊, 徐文静, 宋舒娜, 刘奕君, 余振苏. 人体健康系统科学的建构与实践(口头报告) [C], 第二届中国系统科学大会, 北京, 2018.5.12.
- [3] 张磊, 徐文静, 张莉, 刘奕君, 余振苏. 基于人体量子系统科学的健康管理(口头报告) [C], 第三届中国系统科学大会, 长沙, 2019.5.19.
- [4] 张磊, 宋舒娜, 余钰嵘, 张莉, 李昱洁, 赵前程, 余振苏. 脉诊数理模型辨析人体健康态(口头报告) [C], 第四届中国系统科学大会, 青岛, 2020.9.19.
- [5] 张磊, 宋舒娜, 余钰嵘, 余振苏. 中医智能脉诊研究的新进展(口头报告) [C], 第五届中国系统科学大会, 南京, 2021.5.22.

## 致谢

多年以后，我又想起了第一次听到系统科学的那个遥远的下午。在北科的学术报告厅，座无虚席，很多同学只好搬来凳子，把讲台围得水泄不通。讲台上是北大的马蔼乃教授，一位剪着干净短发的女士，声音沉静有力。她已经七十多岁了，仍然站着讲了三四个小时。报告的题目是《迎接第二次文艺复兴，学习钱学森科学思想》，内容从产业革命到第二次文艺复兴，从现代科学技术体系到社会主义建设总体设计部，从天地人机信息网络到大成智慧教育方案。整个报告大开大合，气势磅礴。台下的我大为震撼，仿佛一个新的天地被打开了。

后来，我跟马老师有了更深的缘分，得到了她许多当面的指导。燕南园 64 号，也成为了我们“钱学森青年学习小组”的主要活动场所。马老师的身教言行，都深深影响着我。在她那里，我还查到了当年报告的 PPT，显示讲座的日期是 2010 年 11 月 28 日。那时，我刚好 20 岁，正是我的黄金时代。当时的我和几个伙伴经常聚在一起，几乎天天去北大蹭课、听讲座。作为一名工科生，我们感兴趣的却都是文史哲的课程，关心的是国家金融经济和国际局势。每天读书看报，中午约定在食堂角落一起吃饭就变成了讨论会，颇有些指点江山，激扬文字的味道。现在回想，那真是人生中最美好的时光之一。但是，我们当时不免看到的更多是存在的问题，而对于如何解决这些问题却看不到出路。

但是，这次讲座给我的感受是前所未有的。我几乎什么都没听懂，但听懂与否已然不重要了。我的话变得很多，很想表达什么，却又不知该说什么。就像是一束光突然打了下来，似乎在说，你就应该去做这个。于是我就开始查北大课表，果然看到一门叫《复杂性科学与哲学》的课程，恰好就是在第二年的春季学期开设。

这门课是由北大哲学系赵光武老师主持的，每一讲都是由清华、北师大、人大等系统科学领域内的专家来讲，从哲学概念讲到方法技术、再到工程实践，已经开设了十多年了。但这样一门前沿的课程，选课的学生却极少，反而前排坐着的都是一群白发苍苍的社会人士。他们每讲必到，风尘仆仆中眼神坚毅，总是在纸上密密麻麻地记着笔记，这种学习精神十分打动人。但老实说，这门课程我仍然听得是云里雾里，脑子里只留下了一些概念，只是觉得很有意思，也认为未来的科学就应该朝这个方向发展，但是如何下手开展研究，或者运用到自己的专业上，却不是很清楚。

所以，一起来听课的同伴，最后也就只剩下我和另一位朋友。他后来也因此认识了余老师，并在当年夏天保送了余老师的研究生。我也因此有更多机会跟余老师交流，没想到后来也成为余老师的博士生。而另一位授课的北师大的狄增如老师，成了我博士

论文答辩委员会的委员之一。这些都是当时的我无法想象到的。人生中的各种缘分，真是非常奇妙。

作为这门课程的“工程实践”部分，余老师的这一讲介绍了许多具体的复杂系统工程的实践。从运动员训练，到中医专家系统，每个话题都十分吸引人。准确的说，这堂课是由倪志勇师兄主讲的，余老师因为有事在下课前才赶到，只讲了结尾。但是他如艺术家一般的气质，激情昂扬的发言，一下子就感染了大家。下课后，许多旁听的同学都围着余老师问问题，余老师却一直强调的是，大家要学习“真本事”。什么才是“真本事”呢？余老师指着身边一位师兄说，像他能够XX天不吃饭就是真本事。这句随口的回答，却成为了我百思不得其解的话头。一直到三年后，我真正实践之后才明白了其中的奥义。

2011年秋季学期，余老师开设了《复杂系统科学导论》课程，我旁听了整个学期的课程。当时课堂的讨论氛围极好，针对几个复杂性话题分小组进行研究讨论。还记得我报名参加的是“理想与动力”小组，把这个话题作为一个严肃的科学问题来对待，进行了丰富的调研，然后定期组织讨论，最后撰写报告。通过这学期的学习，我觉得似乎许多问题在余老师那里是有解的。

之后，我又旁听了每年开设的课程。课程内容也逐渐更加聚焦于大学生在学业、健康等方面的具体复杂事物的分析和处理。对此，余老师逐渐提出了一套完整的工程化的解决方案。我们于是开始实践，实践就会有效果，有了效果自然就会琢磨背后的原因。比如，当时课程从系统优化的角度提倡要“负熵饮食”，即饮食要清淡，少吃多嚼。于是我和几个朋友便开展实践，果然吃得少了，精力反而更加充沛。就这样，因缘际会，之后又进行了更深入的科学实验，让我对人体系统有了全新的认识。一个生机盎然的世界在我眼前诞生了，我看到人作为万物之灵在其中的精妙，于是便立下志向，要对人体系统开展研究。在这个过程中，我也跟余老师建立起了深度的信任。2014年，在我硕士快毕业的时候，收到了余老师让我来新成立的健康系统工程研究所工作的邀请。

当时，我们想做的是一件事是建立“人体模型”，挖掘中医理法技背后的科学内涵，对人体这个复杂系统开展理解，服务于国民的健康保障。2017年，在新奥集团新绎健康研究院的支持下，这一研究逐渐落实在了对中医脉诊的分析建模上。这可以说是完全从头做起，从第一条数据开始测量，从第一行代码开始写起，从第一篇文献开始阅读。这其中的挑战可想而知，但是大家都干劲十足，做得热火朝天。2018年—2019年，我们一边上课、做研究，一边还要每隔几天去灵兰中医馆测量数据。由于每次要测量左右手寸关尺六个部位的脉搏波，还需要填写量表，所以测量一位志愿者至少需要半小时以上。测量期间，志愿者还得尽量保持自然不动的状态。可即使这样，依然有很多来看中医的人们愿意免费参与测量。他们中有的是年轻的外交官，有的是退休的老人，有的是鞋店老板，有的是在校学生……全凭着对于中医的信任，来支持和鼓励我们，希望

看到我们的成果。还有一些多次测量的志愿者，每次都很关心自己的数据情况，希望我们能给出科学的解读，我们也的确从中发现了许多很有意思的信息。测量间隙，我们便在一旁跟诊，看中医师是如何给人把脉，如何诊断治疗。有机会，我们也上手摸脉，然后再跟中医师、机器测量的结果进行对比，体会其中的关系。就这样，在中医临床一线中，看到了中医神奇而又可靠的疗效，也体会到了中医客观化中还存在许多需要解决的问题。寒暑假也成了积累数据的好时机，家人们自然不必说了，还有爸妈的工友，聚会时的同学，来拜年的亲戚，都先来测个脉再说。

这些点点滴滴汇集起来，不知哪一天，竟成了可吞日月的大江大河。在孤独前行的时候，在论文迟迟不能发表的时候，在疫情期间无法进行测量实验的时候，一想到他们，心中便有了再来一次的勇气。哪里有什么困难呢？大不了去睡一觉，从头来过。

所以，如果这篇论文的工作算的上是一点成绩的话，那么这个成绩的获得是跟背后有强大的团队支持分不开的，是跟给予我们帮助的每位人士分不开的。所以，借此机会，我要对他们表示真诚的感谢。

首先，我要感谢我的导师余振苏教授。我们之所以有上面提到的这种不畏困难、勇于创新的勇气，得益于余老师平时的亲身示范。余老师是流体力学界的大师，又对祖国传统文化有几十年的深耕。他有着年青人一般旺盛的活力乃至脾气，但同时又有着非一般的哲学观和方法论、物理直觉和科研品味。他对很多事物总是能够娓娓道来，又能一针见血地亮出其中的问题。他既是高高在上的学界权威，又能带大家一起对酒当歌。总之，这些看似对立的二面集于一身，使得他是一位极具个人魅力的人，总有一种“自信人生二百年”的气质。这并不是因为他没有遇到过挫折，相反，做这些大跨度的创新，肯定要承受更大的风险。但是，余老师依然选择这样做。这种遗世独立的态度，也深深地影响了每一位在他身边的学生。所以，在刚开始做脉诊科研时，我们都有股“初生牛犊不怕虎”的劲头，未想太多这件事的复杂度，也未想过可能的利害得失。这一课题虽然跟身边同学的方向差别很大，甚至可能被人所不理解，但是我们乐在其中。反而是现在回过头再看，对人体系统的复杂度有了更为直观的感受，才明白研究中需要的勇气和品味才是最为可贵的。而这又很难通过刻意练习而获得，只能是在朝夕相处中，潜移默化地从老师们身上继承而来。

行笔至此，我觉得我还要感谢人生道路上其他所有教育过我的老师。很幸运，能够在每个人生阶段，都能遇到非常好的老师。比如，我的小学数学老师徐斯英女士，是一位非常认真和敬业的老师。她总是能够用一半的时间上课，一半的时间让我们当堂完成作业，并且完成批改和改错。所以整个小学期间，我们大部分时间都没有回家写过数学作业，但获得满分的成绩却是最多的。再比如，我的初中班主任王萍老师，她告诉我们，认认真真看几页书，比稀里糊涂翻几本书要好。她的方法很简单，就是把英语书上的每个句子理解了，背下来，然后再一字不差的默写出来。再比如，我的高中班主任黄

学斌老师。高中三年，他每天都是很早就到了班级里，他跟我们讲的最多的，依然是让我们“认真”。还有北大自行车修理铺的蔺师傅，总是一言不发就把车的问题解决了。还有研究生教务李曼老师，能够记得每位同学的基本情况，并把繁琐的毕业流程清晰化，督促和鼓励我们一项项地完成。我要感谢他们在工作中教会我认真和敬业。

我要感谢课题组这个大家庭。感谢陈军老师、毕卫涛老师在平时学习和生活上的关心，陈老师竟通读了本文并做了仔细的批注，毕老师在我试讲时提的许多问题都在正式答辩时被问到了。特别要感谢李荣博士，他在2020年上半年带领我们几个小同学开展了新冠科研的攻关，那又是一段难忘的日子。正是通过这个训练，在下半年我就可以相对独立地上手科研了。感谢宋舒娜博士后带领我们健康组开展科研攻关，从2017年开始，对涉及到方方面面进行设计和安排，非常不容易。我还要特别感谢我们“脉诊攻关”三人组的忻煜和钰蝶。本文工作中的很多顺利或不顺利，都是和他们一起经历和分享的。他们也在帮助我完成博士论文的写作中出力极多，衷心地希望他们能够做出更多出色的成果。还要感谢健康组的奕君，前程，光会，国婷，睿麒，景和，毅卓等同学，你们为健康家园做出了许多贡献。还要感谢湍流组的坦坦、斯雅、韦宙、克馨、超博等同学，你们给实验室带来了很多欢乐。还要感谢湖南中医药大学的晏老师、文老师、高远等，两个团队有着一拍即合的默契。还要感谢虽然已经毕业但曾给与很多帮助的倪志勇师兄、陈曦师兄、薛勇师兄、俞燕明师兄、王莹姐、唐帆、文静、舒晴、泽夏、君艇、师尧、文丰、永煌、余凯、昱洁等，他们都曾贡献过宝贵的数据和研讨。最后，祝愿大家在未来的事业中，各尽其能，各得其所。

我还要感谢为本研究提供了极大帮助的单位和人士。感谢新奥集团新绎健康研究院，为本文的脉诊研究提供项目支持和硬件支持。感谢北京市灵兰中医馆、北医三院二门诊、畅春园社区，为本文采集脉搏波数据提供支持。感谢北大中医学社的指导老师陈璞医师的指导。感谢北大中医学社创社社长王正山医师在多方面的帮助和指导。感谢每一位参与脉诊测量的志愿者。特别感谢在传统文化学习方面给我指导的多位明师和朋友。

我还要感谢“神奇四侠”，正是因为有你们，生命中有了许多闪亮的日子。感谢勇哥，最初脉诊数据处理的程序框架，就是由他完成的。在毕业后，也依然经常关心我的科研进展，经常能给我一些重要的启发。感谢梦娟，给我们带来很多欢声笑语。感谢我的妹妹张莉，感谢生命中有你。

我还要感谢我的父母。他们都是普通人，艰难困苦的环境造就了他们善良勤奋和吃苦耐劳的精神。他们虽然不太知道我在做什么，但仅仅凭着“读书就是好的”这朴素的认识，一直鼓励支持着我。正是因为他们身上的这些品质，我无论走到哪里，依然能够感受到平凡中所蕴蓄的伟大。我还要感谢我的岳父岳母，感谢他们对于我们小家的关心和照顾，让我们始终能以一种轻松的状态面对发生的各种挑战。

我特别要感谢何兮博士。感谢她日常的督促与鼓励，也感谢她为我的科研提供了丰富的数据。前天是我们在一起的第五千天。可不知从何时起，当初贪玩的少女，竟可以在疫情严重的时候一人闭关三个月，也可以将哲学大部头一本接一本地看过去，还可以几年如一日地做一件很小的事情。光阴不负，不负光阴，感谢一路相伴。

最后，我要感谢北京大学。正是因为北大强烈的民族使命感，兼容并包的学术氛围，才会允许有本文这样大跨度的创新性研究。世人张目望，北大当自强。燕园湖水长养着这里的每一个人，造就了我们独特的精神气质。在努力创造一个更加美好的世界的道路中，我们仍要继续努力。



## 北京大学学位论文原创性声明和使用授权说明

### 北京大学学位论文原创性声明和使用授权说明

#### 原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

论文作者签名： 日期： 年 月 日

#### 学位论文使用授权说明

(必须装订在提交学校图书馆的印刷本)

本人完全了解北京大学关于收集、保存、使用学位论文的规定，即：

- 按照学校要求提交学位论文的印刷本和电子版本；
- 学校有权保存学位论文的印刷本和电子版，并提供目录检索与阅览服务，在校园网上提供服务；
- 学校可以采用影印、缩印、数字化或其它复制手段保存论文；
- 因某种特殊原因需要延迟发布学位论文电子版，授权学校一年/两年三年以后，在校园网上全文发布。

(保密论文在解密后遵守此规定)

论文作者签名： 导师签名：

日期： 年 月 日





## 学位论文答辩委员会名单

### 学位论文答辩委员会名单

<b>论文题目</b>	应用于新冠疫情演化和脉搏波分析的人体序参量理论研究			
<b>作 者</b>	张磊			
<b>专 业</b>	力学（生物力学与医学工程）			
<b>答辩委员会成员</b>	<b>姓 名</b>	<b>专业技术职称</b>	<b>从事专业</b>	<b>工作单位</b>
主席	谭文长	教授	力学（生物力学与医学工程）	北京大学
委员	狄增如	教授	系统理论	北京师范大学
	荣起国	副教授	力学(生物力学与医学工程)	北京大学
	熊春阳	教授	力学（生物力学与医学工程）	北京大学
	晏峻峰	教授	中医	湖南中医药大学



## 博士学位论文答辩委员会决议书

### 北京大学博士学位论文答辩委员会决议书

(本表由博士答辩委员会秘书填写,一式两份,一份存学校档案,一份存研究生个人档案)

院、系: 工学院

专 业: 力学(生物力学与医学工程)

姓 名: 张磊

研究方向: 生物医学信息技术

学 号: 18011111654

导师姓名: 余振苏

**学位论文题目:** 应用于新冠疫情演化和脉搏波分析的人体序参量理论研究

答辩委员会对学位论文和答辩情况的学术评语 (主要就论文选题意义, 创新性成果及学术水平; 论文存在的主要不足之处, 以及博士生答辩情况等方面)。

本文将人体序参量理论应用于分析疫情演化和人体脉搏波分析, 对于认识疫情发展的规律和人体个性化的健康辨析, 具有理论意义和应用价值。论文的主要创新成果如下:

第一, 发展了一套人体复杂系统研究的新范式。该范式从对称性分析开始, 通过选择能够刻画人体系统有序度的序参量, 来开展亚类分解和定量建模, 为认识人体复杂系统自组织机制, 突破传统统计平均方法的局限, 开展人体个性化特征的数学建模提供了一种新工具。

第二, 应用上述范式于新冠疫情分析, 构建了GHE演化动力学新模型, 成功地对社会抗疫力度开展了定量描述, 定量刻画了全国不同省市、不同流行波次的社会抗疫效果。在应用于全球52个国家的死亡率研究中, 定量解释了各国2到3个数量级差别的死亡率, 同时发现各国早期非存活者在ICU的持续时间与当地疫情持续时间之间的普适线性律。

第三, 应用上述范式于人体脉波分析, 创新性地提出了基于“中医脉诊相空间”来划分亚类的方法, 并根据中医“知常达变”理论构建了“偏离度”这一关键序参量, 实现了对人体健康态的个性化和日变健康指数的刻画。

该论文方法创新、内容丰富、数据可靠、调研充分、论述清楚、分析严谨、层次分明、写作规范。论文工作表明, 作者掌握了本学科领域坚实宽广的基础理论和系统深入的专门知识, 具有独立从事科学研究的能力。答辩过程中, 表述清楚, 回答问题准确。

答辩委员会经无记名投票表决, 一致同意张磊通过博士论文答辩, 建议授予张磊工学博士学位。

答辩委员会表决结果: 实到答辩委员: 5 人,

同意博士毕业者 5 人, 不同意博士毕业者 0 人, 结论为 准予毕业

同意建议授予博士学位者 5 人, 不同意建议授予博士学位者 0 人, 结论为 建议授予学位

答辩时间: 2022年06月04日9时

答辩委员会主席(签字):

谭文长

委员(签字):

黄明 袁峰光 孙敬峰 熊春阳



本表不得折叠、不得跨页



## 提交终版学位论文承诺书

### 提交终版学位论文承诺书

本人郑重承诺，所呈交的学位论文为最终版学位论文。本人知晓，该版学位论文将用于校学位评定委员会审议学位、国家和北京市学位论文抽检。论文版本呈交错误带来的结果将由本人承担。

论文作者签名: 张名石

日期: 年 月 日

