

Article

Artificial Intelligence in Postoperative Care: Assessing Large Language Models for Patient Recommendations in Plastic Surgery

Cesar A. Gomez-Cabello ¹, Sahar Borna ¹, Sophia M. Pressman ¹, Syed Ali Haider ¹, Ajai Sehgal ²,
Bradley C. Leibovich ^{2,3} and Antonio J. Forte ^{1,2,*}

¹ Division of Plastic Surgery, Mayo Clinic, Jacksonville, FL 32224, USA

² Center for Digital Health, Mayo Clinic, Rochester, MN 55905, USA

³ Department of Urology, Mayo Clinic, Rochester, MN 55905, USA

* Correspondence: forte.antonio@mayo.edu

Abstract: Since their release, the medical community has been actively exploring large language models' (LLMs) capabilities, which show promise in providing accurate medical knowledge. One potential application is as a patient resource. This study analyzes and compares the ability of the currently available LLMs, ChatGPT-3.5, GPT-4, and Gemini, to provide postoperative care recommendations to plastic surgery patients. We presented each model with 32 questions addressing common patient concerns after surgical cosmetic procedures and evaluated the medical accuracy, readability, understandability, and actionability of the models' responses. The three LLMs provided equally accurate information, with GPT-3.5 averaging the highest on the Likert scale (LS) (4.18 ± 0.93) ($p = 0.849$), while Gemini provided significantly more readable ($p = 0.001$) and understandable responses ($p = 0.014$; $p = 0.001$). There was no difference in the actionability of the models' responses ($p = 0.830$). Although LLMs have shown their potential as adjunctive tools in postoperative patient care, further refinement and research are imperative to enable their evolution into comprehensive standalone resources.



Citation: Gomez-Cabello, C.A.; Borna, S.; Pressman, S.M.; Haider, S.A.; Sehgal, A.; Leibovich, B.C.; Forte, A.J. Artificial Intelligence in Postoperative Care: Assessing Large Language Models for Patient Recommendations in Plastic Surgery. *Healthcare* **2024**, *12*, 1083. <https://doi.org/10.3390/healthcare12111083>

Academic Editor: Joaquim Carreras

Received: 11 April 2024

Revised: 13 May 2024

Accepted: 23 May 2024

Published: 24 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: large language models; artificial intelligence; plastic surgery; postoperative care; patient resource; patient-centered outcomes; patient satisfaction

1. Introduction

The advancement of artificial intelligence (AI) offers new opportunities for healthcare improvements and individualized patient care. Large language models (LLMs) represent a breakthrough in applied AI and medical practice. They learn and understand the complex patterns and structures in typical language by leveraging natural language processing (NLP) and deep learning (DL) techniques, in particular, transformer architectures [1]. LLMs can process, interpret, and summarize vast amounts of internet data in real time and generate human-like text responses [2,3]. Since their release, LLMs have shown promise in providing accurate medical knowledge throughout distinct medical fields, encouraging the medical community to actively explore their potential applications and determine how to best leverage their capabilities.

Plastic surgery is a constantly innovating field that relies on updated, accurate tools to provide patient-centered outcomes [4]. In particular, cosmetic surgery is a major ever-evolving field with a rising demand due to its growing social acceptance. In 2022, there were 26.2 million surgical and minimally invasive cosmetic and reconstructive procedures in the United States, with a 19% increase in cosmetic surgery since 2019 [5]. Plastic surgeons are challenged to provide comprehensive information and support to an increasing number of patients [3]. LLMs, in particular, Open AI's ChatGPT, have proved to be helpful throughout the specialty, with the ability to generate and detect areas for improvement and potential

research [4]. ChatGPT has even proved to have a comparable level of knowledge to 50% of first-year integrated plastic surgery residents, as shown by in-service exam scores [6].

While most LLM research focuses on their clinical decision capabilities, such as diagnosis and treatment, one very promising application is as a patient resource. Given its accessibility, patients already rely on the Internet as their primary source of information and make decisions based on the information they find, risking themselves being misinformed [2,7]. Previous studies have shown that despite some inherent limitations, LLMs can provide accurate information for patients seeking insights into cosmetic procedure details, outcomes, risks, and benefits [2,3,8]. Additionally, they offer immediate, multilingual, round-the-clock access to information, which is crucial for addressing concerns and questions outside of regular office hours and ensuring equity.

In plastic surgery, the postoperative period is as important as the surgical procedure itself, and in some cases even more so, as the patient's final outcome and the procedure's long-term success depend on it [9]. Although surgeons offer clear verbal and printed postoperative care recommendations, patients may struggle to retain or implement them due to their postoperative emotional state. This forces them to keep looking for their surgeons and their staff's help outside of office hours, sometimes overburdening the surgical staff. Approaches to decrease these burdens may include limiting notifications or delegating responses to less-trained support staff [10]. These limit access to high-quality care and lead to the common patient misconception that they are forgotten once their surgery has finished. The immediate and accurate responses provided by LLMs can reduce anxiety and improve patient satisfaction [11–13].

Building on previous work assessing LLMs' use in the perioperative period [2,3,8,14,15], we aimed to evaluate and compare the ability of the currently readily available LLMs, ChatGPT-3.5, ChatGPT-4, and Google's Gemini, formerly Bard, to provide postoperative care recommendations to patients who underwent any one of the five most common cosmetic plastic surgery procedures according to the American Society of Plastic Surgeons (ASPS) in 2022 [5] in the absence of utilizing a retrieval-augmented generation (RAG) approach. As accessibility is fundamental to achieving equal health distribution, we decided to include the publicly available version of ChatGPT in our comparison.

2. Methods

2.1. Study Design

We created a set of 32 questions addressing the most common concerns patients have at our clinic after liposuction ($n = 7$), breast augmentation ($n = 6$), abdominoplasty ($n = 6$), mastopexy ($n = 6$), and blepharoplasty ($n = 7$). Each question was presented to each LLM using a combination of specialized medical terminology and common language to reflect the various ways in which a typical plastic surgery patient might phrase their questions. Moreover, we did not employ any specialized prompting engineering techniques, such as contextualization or role play, to ensure that the responses obtained were not a result of these methods, as a typical patient may not be aware of them or be comfortably familiarized with them. We only asked each question once and in different chats. After every question was asked to one model, another was tested. Figure 1 portrays an example of the questions. In addition, we have included the full set of questions provided to the LLMs and the responses retrieved in Supplementary File S1 and in a Supplementary File S2, respectively.

Question	GPT-3.5	GPT-4	Gemini
I had a breast lift some years ago and I feel that my breast are starting to sag, is that something normal or should I go with my surgeon?	"It's not uncommon for the results of a breast lift surgery (mastopexy) to change over time. Factors such as aging, gravity, pregnancy, weight fluctuations, and genetics can all influence breast shape and position. It's essential to understand that breast tissue can naturally sag or droop over the years, even after a successful breast lift procedure. However, whether or not you should consult with your surgeon depends on your specific concerns and goals. Here are some considerations: Normal Aging Process: Over time, the effects of gravity, loss of skin elasticity, and changes in breast tissue can lead to a gradual return of some degree of sagging. This is a normal part of the aging process. Post-Surgery Changes: While a breast lift can provide long-lasting results, it does not stop the aging process. Your breasts may continue to change in appearance over the years, which could include some degree of sagging. Consultation with Your Surgeon: If you are concerned about changes in the appearance of your breasts or feel that they are sagging more than you are comfortable with, it's a good idea to consult with your original surgeon or a board-certified plastic surgeon. They can evaluate your current breast condition and discuss potential options for improvement. Breast Lift Revision: In some cases, a breast lift revision surgery may be recommended if you are dissatisfied with the degree of sagging. A revision procedure can help restore breast shape and position. Your surgeon will assess your specific needs and discuss the best approach for achieving your desired outcome. Lifestyle Factors: Lifestyle factors such as maintaining a stable weight, wearing supportive bras, and protecting your skin from excessive sun exposure can help minimize the progression of breast sagging over time. Realistic Expectations: It's important to have realistic expectations about the results of cosmetic surgery. While a breast lift can provide significant improvement, it may not completely halt the natural aging process. Consulting with your surgeon is the first step in addressing your concerns and exploring potential solutions. They can provide personalized advice and recommendations based on your individual situation and goals."	"It's not uncommon for some degree of breast sagging to occur after a breast lift, especially as the years pass. This can be due to several factors, including aging, gravity, changes in weight, hormonal fluctuations, and the natural loss of skin elasticity over time. Even after surgical lifting, these factors can contribute to changes in breast shape and position. However, if you notice significant changes or if you're concerned about the appearance of your breasts, it's a good idea to consult with your surgeon. They can assess whether what you're experiencing is within the normal range of changes following a breast lift or if there might be a need for a revision surgery or other interventions. Your surgeon can provide personalized advice based on your specific situation, including your initial surgical outcome, your current concerns, and your overall health and goals. Maintaining a stable weight, wearing supportive bras, especially during exercise, and taking care of your skin can help maintain your results for as long as possible. But, it's important to have realistic expectations and understand that some changes over time are inevitable. If you're feeling unsure or worried, scheduling a consultation with your surgeon is the best way to get peace of mind and make informed decisions about your care."	"It's important to understand that the results of any cosmetic procedure, including breast lifts, are not permanent. While they can offer significant improvements in appearance and satisfaction for many years, various factors can contribute to a return of sagging over time. These can include: Natural aging: As we age, our skin loses elasticity and collagen, which can lead to drooping of tissues throughout the body, including the breasts. Gravity: Gravity continuously exerts a downward pull on the breasts, and while a lift addresses this to some extent, it can't completely counteract its effects over time. Weight fluctuations: Significant weight gain and loss can stretch the skin and underlying supporting tissues, contributing to sagging. Pregnancy: Subsequent pregnancies after a lift can stretch the skin and affect the results. Given that it's been several years since your procedure and you're experiencing renewed sagging, it's not necessarily unusual but consulting with your original surgeon is highly recommended. They are best equipped to: Evaluate your specific situation: They can assess the degree of sagging, the condition of your skin, and any other contributing factors. Discuss your options: Based on their evaluation, they can discuss whether additional surgery might be appropriate, or if alternative methods like supportive bras or specific exercises might be beneficial. Address your concerns: They can answer your questions, address any anxieties you might have, and provide you with realistic expectations for potential next steps. Remember, maintaining open communication with your original surgeon throughout the process is crucial for achieving the best possible outcome and addressing any concerns you may have."
What can I do to avoid or decrease edema after eyelid surgery?	"Edema, or swelling, is a common occurrence after eyelid surgery (blepharoplasty). While some degree of swelling is normal during the initial stages of recovery, there are steps you can take to minimize and manage it. Here are some tips to help avoid or decrease edema after eyelid surgery: Follow Your Surgeon's Instructions: Adhere to all post-operative instructions provided by your plastic surgeon. These instructions are tailored to your specific case and are essential for a smooth recovery. Elevate Your Head: Sleeping with your head elevated at a 30-45 degree angle can help reduce swelling. Use extra pillows to prop yourself up, or consider sleeping in a reclining chair for the first few days. Cool Compresses: Applying cold compresses to the eyelids can help reduce swelling and discomfort. Use clean, cold, damp washcloths or gel packs, and follow your surgeon's recommendations for timing and duration. Avoid Heat: Avoid hot showers, hot tubs, saunas, or anything that increases blood flow to the face during the early stages of recovery, as heat can exacerbate swelling. Limit Physical Activity: Rest and avoid strenuous physical activities in the first few days to minimize swelling. Gradually increase your activity level as advised by your surgeon. Hydrate: Drink plenty of water to stay hydrated, as proper hydration can help reduce swelling and promote healing. Eye Drops: If prescribed or recommended by your surgeon, use lubricating eye drops or artificial tears to keep your eyes moist and comfortable. Dry eyes can contribute to discomfort and swelling. Avoid Salt: Reduce your salt intake, as excess sodium can lead to fluid retention and increased swelling. Protect Your Eyes: Use any protective measures, such as eye shields or glasses, as instructed by your surgeon to prevent accidental rubbing or pressure on the eyes. Avoid Allergens: Minimize exposure to allergens, such as dust and pollen, that can irritate the eyes and worsen swelling. Avoid Alcohol and Caffeine: Alcohol and caffeine can dehydrate the body, so it's best to limit or avoid them during the early recovery period. Proper Nutrition: Follow a well-balanced diet rich in fruits and vegetables, which can support healing and reduce inflammation. Stay Relaxed: Stress and anxiety can contribute to swelling, so try to stay relaxed and avoid situations that may cause unnecessary stress. Follow-Up Appointments: Attend all follow-up appointments with your surgeon so they can monitor your progress and make any necessary adjustments to your post-operative care plan. It's important to be patient during the recovery process, as swelling can take time to subside completely. In most cases, the majority of swelling should resolve within a few weeks, with gradual improvement over the following months. If you have concerns about the extent of swelling or any other post-operative issues, contact your surgeon for guidance and reassurance"	"Post-operative edema (swelling) is a common occurrence after eyelid surgery (blepharoplasty) and typically resolves over time. However, there are several strategies you can employ to minimize swelling and promote a smoother recovery. Here are some tips: Cold Compresses: Applying cold compresses or ice packs to the eye area can significantly reduce swelling, especially during the first 48 hours after surgery. Be sure to wrap the ice pack in a clean cloth to avoid direct contact with your skin. Elevation: Keep your head elevated higher than your chest when lying down. This can be achieved by using extra pillows under your head while sleeping or resting. Elevation helps reduce swelling by improving venous return and decreasing fluid accumulation. Avoid Strenuous Activities: Limit activities that can increase blood pressure to your head, such as bending over, lifting heavy objects, or performing strenuous exercises, as they can exacerbate swelling. Follow Your Surgeon's Instructions: Your surgeon will provide specific aftercare instructions, including how to care for your incisions, when to return for follow-up visits, and the use of any prescribed medications or ointments. Adhering to these instructions is crucial for minimizing swelling and preventing complications. Stay Hydrated: Drinking plenty of water can help reduce swelling by flushing excess fluids out of your body. Limit Salt Intake: Excessive salt can contribute to fluid retention. Reducing your salt intake can help minimize swelling. Avoid Smoking and Alcohol: Smoking can delay healing, and alcohol can lead to fluid retention. Both should be avoided during the recovery period. Use Prescribed Medications: If your surgeon prescribes medications to help reduce swelling, such as anti-inflammatory drugs, be sure to take them as directed. Protect Your Skin: Follow any recommendations for sun protection, as sun exposure can contribute to swelling and affect healing. Healthy Diet: Eating a balanced diet rich in vitamins and minerals can support your body's healing process. Remember, recovery times and experiences vary from person to person. If you have concerns about swelling or any aspect of your recovery, it's important to contact your surgeon. They can provide advice tailored to your specific situation and intervene if necessary to address any complications."	"While completely avoiding edema (swelling) after eyelid surgery is not possible, there are several steps you can take to minimize its severity and duration: Cold compresses: Apply cold compresses to your eyelids for 10-15 minutes at a time, several times a day, especially during the first 24-48 hours after surgery. Wrap a gel ice pack or a clean washcloth soaked in cold water in a thin towel to prevent direct contact with the skin. Head elevation: Keep your head elevated at night by using extra pillows or sleeping in a reclining chair. This helps reduce fluid accumulation around the eyes. Avoid straining: Avoid strenuous activities, bending over, or lifting heavy objects for at least 2 weeks after surgery. These activities can increase blood pressure in your head and worsen swelling. Limit salt intake: Reducing your salt intake can help minimize fluid retention throughout your body, including around your eyes. Quit smoking: Smoking can impair healing and increase the risk of complications, including swelling. If you smoke, quitting at least 4 weeks before surgery is recommended. Gentle cleansing: Avoid rubbing your eyes or using harsh soaps or cosmetics around the eyes. Gently cleanse the area with a mild cleanser as instructed by your doctor. Sleep: Getting enough sleep is crucial for healing. Aim for 7-8 hours of sleep each night. Medications: Your doctor may prescribe pain medication or anti-inflammatory medication to help manage swelling and discomfort. Take these medications exactly as directed. Remember, these are general recommendations, and it's crucial to follow your doctor's specific instructions for post-operative care after eyelid surgery. They will provide you with detailed guidance on how to care for your eyes and minimize swelling."

Figure 1. Cont.

Question	GPT-3.5	GPT-4	Gemini
I just had a liposuction and I feel like I am gaining weight, is this normal?	"I'm not a doctor, but I can try to provide you with some general information about liposuction and weight fluctuations. Liposuction is a surgical procedure designed to remove fat cells from specific areas of the body. While it can result in immediate reductions in localized fat, it's important to understand that liposuction doesn't prevent future weight gain. If you're experiencing weight fluctuations after liposuction, several factors could be contributing to this: Swelling: Swelling is a common side effect after liposuction, and it can take several weeks to months for the swelling to subside completely. During this period, you may feel like you've gained weight due to the increased fluid retention and inflammation. Lifestyle and Diet: Your post-liposuction weight can be influenced by your lifestyle and dietary choices. If you're consuming more calories than you're burning, you can still gain weight, even after the procedure. Redistribution of Fat: Liposuction removes fat from specific areas, but it doesn't affect fat distribution in the rest of your body. This means that if you consume excess calories, your body may store fat in other areas, potentially leading to weight gain. Muscle Changes: After liposuction, it's important to maintain or build muscle tone in the treated areas to prevent sagging and maintain a balanced appearance. If you're not engaging in regular exercise, muscle loss can contribute to changes in your body's shape. Psychological Factors: Sometimes, people may experience changes in their body image or perception after cosmetic procedures, which can lead to concerns about weight gain even if there's no significant physical change. If you have concerns about weight gain or changes in your body after liposuction, it's crucial to consult with your plastic surgeon or a medical professional who can provide personalized advice and address your specific situation. They can help you understand the causes of any changes you're experiencing and provide guidance on maintaining a healthy lifestyle to achieve your desired results."	"After liposuction, experiencing fluctuations in weight is not uncommon. Here are a few points to consider that might explain why you feel like you're gaining weight: Swelling: The most common reason for feeling like you've gained weight shortly after liposuction is swelling. The body's response to the surgery involves retaining fluid in the treated areas, which can make you feel heavier or look puffier than before. This swelling can last for several weeks to months. Fluid Retention: In addition to swelling in the surgical area, your body might retain fluids more generally as part of the healing process. This can also contribute to a sensation of weight gain. Diet and Activity Levels: Your diet and activity levels post-surgery can also impact your weight. If you're consuming more calories than you're expending, especially during your recovery period when your physical activity might be limited, you might notice weight gain. Temporary Weight Gain: It's also worth noting that liposuction removes fat cells from the body, but if you don't maintain a stable weight by balancing calorie intake with physical activity, the remaining fat cells in your body can grow, leading to weight gain. Healing Process: The body is in a state of healing and recovery after liposuction, and it can take time for the final results to be visible. The initial weight gain or bloated feeling usually resolves as the body heals. Lymphatic System Response: The lymphatic system plays a crucial role in removing fluids and waste from your body. Liposuction can temporarily disrupt the normal function of the lymphatic system, leading to fluid accumulation. It's important to follow your surgeon's post-operative instructions closely, including wearing compression garments if recommended, maintaining a healthy diet, staying hydrated, and gradually returning to exercise under guidance. These steps can help minimize swelling and support your recovery. If you're concerned about your weight gain after liposuction, it's best to discuss this with your surgeon. They can assess your situation to determine whether what you're experiencing is a normal part of the healing process or if there's another underlying issue that needs to be addressed."	"While it's understandable to feel concerned about weight gain after liposuction, it's important to remember that experiencing a perceived increase in weight after the procedure can be normal. Here's a breakdown of the points to consider: - Liposuction is not a weight loss solution: Liposuction aims to reshape specific areas by removing fat cells. It's not intended for overall weight loss. - Initial weight fluctuations: * Fluid retention: During healing, your body might retain fluids, leading to a temporary increase on the scale. * Inflammation: The surgical process itself can cause inflammation, which can contribute to a feeling of puffiness or weight gain. - Fat cells vs. remaining fat: * The treated areas have fewer fat cells, so even with weight gain, they won't expand as much as untreated areas. * However, the remaining fat cells throughout your body can still grow if you gain weight overall. Here's what you can do: Consult your doctor: They can assess your individual situation, provide reassurance, and address any specific concerns you have. Maintain a healthy lifestyle: Continue a balanced diet and regular exercise routine to promote overall health and maintain your results. Be patient: Allow your body time to heal fully, which can take several weeks or even months. The final results of liposuction might not be fully visible until several months after the procedure. If you have any significant concerns or questions about your recovery process, always consult your doctor for personalized guidance."
It's been 2 weeks since my breast augmentation surgery, can I start doing yoga?	"After breast augmentation surgery, it's crucial to follow your surgeon's specific post-operative instructions, including guidelines for physical activity and exercise. While I can provide some general information, it's essential to consult with your plastic surgeon before resuming any exercise routine, including yoga. Here are some general considerations for resuming yoga or other forms of physical activity after breast augmentation: Consult Your Surgeon: Your surgeon is the best source of guidance for when it's safe to resume exercise. They will consider your individual healing progress and recommend an appropriate timeline. Early Post-Operative Period (First Few Weeks): During the first few weeks after breast augmentation, you will likely be advised to avoid strenuous activities, including yoga, that involve chest expansion, stretching, or intense movements. Your body needs time to heal, and strenuous exercise can increase the risk of complications. Gradual Return to Exercise: As you progress through your recovery, your surgeon may gradually allow you to reintroduce gentle physical activities. This timeline can vary based on individual factors and the extent of your surgery. Listen to Your Body: Pay close attention to how your body feels during and after yoga or any physical activity. If you experience pain, discomfort, or unusual sensations, stop immediately and consult with your surgeon. Modify Your Practice: When you are cleared to resume yoga, consider modifying your practice to avoid putting excess strain on your chest muscles and incisions. Use props, avoid deep stretches, and choose gentle yoga poses that don't involve significant chest or arm movement. Supportive Bra: Ensure that you are wearing a supportive sports bra or the recommended post-surgical bra during any physical activity to provide adequate support and minimize breast movement. Remember that everyone's recovery timeline is unique, and it's essential to prioritize your healing during this period. Rushing back into strenuous exercise too soon can lead to complications and affect your surgical results. Always consult with your plastic surgeon to ensure that your exercise plan aligns with your recovery and healing process."	"It's important to follow the guidance of your surgeon or healthcare provider when resuming physical activities after breast augmentation surgery. The timeline for when you can start doing yoga varies depending on the individual, the specifics of the surgery, and the type of yoga you plan to practice. Generally, light activities can be resumed within a few days to a week after surgery, but it's crucial to avoid any exercises that involve heavy lifting, strenuous activities, or significant stretching of the chest muscles for at least 4-6 weeks. Some gentle yoga poses that do not put pressure on the chest or involve stretching the arms overhead might be permissible earlier, but poses that strain the chest, back, or shoulders should be avoided until your surgeon gives you the green light. Here are a few guidelines for resuming yoga after breast augmentation: Wait for Clearance: Get explicit clearance from your surgeon before starting yoga or any form of exercise. They might provide specific guidelines based on your recovery progress. Start Slow: Begin with gentle, low-impact poses and avoid any asanas that put pressure on your chest or involve stretching or compressing the surgical area. Listen to Your Body: Pay close attention to your body's signals. If you feel any pain or discomfort, stop immediately and consult your healthcare provider. Modify Poses: Consider modifying poses to reduce strain on your chest and incisions. A yoga instructor familiar with post-surgery recovery can offer personalized modifications. It's essential to prioritize healing and allow your body enough time to recover fully before returning to your regular yoga practice or any physical activities. Recovery times can vary, and rushing the process may lead to complications or prolong the healing period. Always consult with your healthcare provider for advice tailored to your specific situation and recovery progress."	"While it's tempting to get back to your yoga routine after 2 weeks, it's generally not recommended to resume yoga practice so soon after breast augmentation surgery. Most surgeons advise waiting for 4 to 6 weeks for your body to heal adequately before engaging in any activity that puts strain on the chest area, including yoga. Here's why waiting is important: Healing: The initial weeks are crucial for proper healing and minimizing the risk of complications like displaced implants or delayed wound closure. Engaging in yoga poses, especially those involving upper body movements, can put undue stress on the surgical site and hinder the healing process. Discomfort: You might still experience soreness and discomfort in your chest and upper body during the first few weeks. Attempting yoga poses could exacerbate this discomfort and hinder your ability to focus on proper form and alignment. However, it's important to consult with your plastic surgeon for specific guidance. They are best equipped to assess your individual recovery progress and advise on when it's safe to resume specific activities like yoga. They may even suggest modified yoga poses or alternative exercises suitable for your current stage of recovery. In the meantime, you can focus on low-impact activities like walking or gentle stretching to stay active during your recovery."
Is it normal to feel abdominal numbness after a tummy tuck?	"Yes, it is normal to experience abdominal numbness after a tummy tuck (abdominoplasty) procedure. Numbness in the abdominal area is a common side effect of the surgery and is typically due to the disruption of sensory nerves during the procedure. This numbness can extend to various areas of the abdominal region and may vary in intensity from person to person. Here are some key points to keep in mind regarding abdominal numbness after a tummy tuck: Temporary Sensation Changes: Numbness or altered sensation in the abdominal area is usually temporary. Over time, sensation gradually returns, but it can take several months to a year or more for full sensation to come back in some cases. Nerve Regeneration: As your body heals, damaged nerves will begin to regenerate, leading to a gradual return of sensation. However, the process can be slow and may not result in the same level of sensation as before the surgery. Patience is Key: Be patient and allow your body time to heal. Avoid any activities that could put excessive strain on the abdominal area during the initial recovery period to minimize the risk of complications or prolonged numbness. Follow-Up Appointments: Attend all scheduled follow-up appointments with your plastic surgeon. They will monitor your healing progress and can provide guidance on managing any concerns related to sensation changes. Scar Management: Proper scar care and management can also play a role in the overall healing process. Follow your surgeon's recommendations for scar care to optimize the appearance of your tummy tuck scar. It's important to communicate any concerns or questions about sensation changes with your plastic surgeon during your post-operative follow-up visits. They can provide you with personalized advice and address any specific issues related to your recovery. In most cases, abdominal numbness is a normal part of the healing process and should gradually improve over time"	"Yes, it is normal to feel abdominal numbness after a tummy tuck (abdominoplasty). This surgical procedure involves making incisions, tightening muscles, and removing excess skin and fat from the abdomen, which can affect the nerves in the area. Nerve endings in the skin and subcutaneous tissues may be disrupted or damaged during the procedure, leading to temporary or sometimes prolonged numbness in the treated area. The extent and duration of the numbness can vary depending on the specifics of the surgery, such as the amount of tissue removed, the technique used, and individual healing processes. For most patients, the sensation gradually returns over months as the nerves regenerate, but it can take up to a year or more for full sensation to return. In some cases, there may be areas of permanent numbness, but this is less common. It's important to follow your surgeon's post-operative instructions carefully and attend all follow-up appointments to monitor your healing process. If you have concerns about the numbness or if it is accompanied by signs of infection or other complications, you should contact your surgeon for advice."	"Yes, experiencing abdominal numbness after a tummy tuck is completely normal. In fact, it's one of the most common side effects of the surgery. During a tummy tuck, some nerves in the treated area are inevitably disturbed or severed. This disrupts the communication between the nerves and the brain, leading to a temporary loss of sensation in the affected area. The numbness can range from mild to complete and can affect the entire abdomen, specific areas like the lower abdomen or around the scar, or even extend to the upper thighs. You might also experience tingling, prickling, or a feeling of tightness in the area. The good news is that this numbness is usually temporary. In most cases, sensation gradually returns over time, with most patients regaining normal feeling within 6 to 9 months after surgery. However, it's important to note that some patients may experience a permanent slight decrease in sensation, particularly around the scar. Here are some additional things to keep in mind: It's important to be patient: Don't expect your feeling to return overnight. The healing process takes time, and nerve regeneration can be slow. Talk to your doctor: If you are concerned about the numbness or if it seems to be worsening, be sure to speak to your doctor. They can assess your individual situation and provide guidance. Gentle massage: Once your incisions are fully healed and your doctor gives the go-ahead, gentle massage of the affected area may help promote nerve healing. If you are considering a tummy tuck, it is important to be aware of the potential for temporary numbness and to discuss this with your doctor beforehand."

Figure 1. Examples of the questions provided to the LLMs.

2.2. Employment of Language Models

Our goal was to evaluate the capabilities of the currently available LLMs to the public without utilizing an RAG approach. To do this, we used the models' responses based on their current training. After learning grammar, vocabulary, and context during pre-training with a vast dataset of internet text, the models are fine-tuned on specific datasets tailored to specialized tasks such as text generation or conversation [1]. Models such as GPT and Gemini are further trained using extensive internet-sourced text data, such as books, articles, wikis, and websites, including high-level-evidence medical research [1,14,16]. This training enables the models to provide accurate medical information.

2.3. Evaluation Tools

We evaluated and compared each LLM response's accuracy, readability, understandability, and actionability.

For medical accuracy, we utilized a 5-point Likert scale with the following values: 1 point: completely incorrect, 2 points: partially incorrect, 3 points: partially correct and incorrect, 4 points: partially correct, and 5 points: completely correct. To score each answer, we used as ground truth the ASPS's webpage [17–21] and textbooks such as *The Art of Aesthetic Surgery* [22] and *Essentials of Plastic Surgery* [23]. Three independent authors (C.A.G.C., S.B., and S.A.H.) analyzed and graded the responses. Any discrepancies were discussed to reach a consensus, and when not possible, the most common score assigned by the authors was used.

We used the Flesch Reading Ease (FRE) score and the Flesch–Kincaid Grade Level (FKGL) to assess readability. The FRE gives a score between 1 and 100, with scores around 100 meaning that the document is extremely easy to read. Scoring between 70 and 80 is equivalent to school-grade level 8. The FKGL assesses the approximate reading grade level of a text. If a text has an FKGL of 8, the reader needs a grade 8 reading level or above to understand it. Both tests take into account the number of sentences, words, and syllables to emit a score [24]. According to the American Medical Association (AMA) [25] and the National Institute of Health (NIH) [26], readability scores should not exceed 6th and 8th-grade levels, respectively. We calculated the FRE score and the FKGL for every model's response using a free online calculator.

We employed the Patient Education Materials Assessment Tool (PEMAT) [27] to measure the understandability and actionability of the LLMs' responses. While it is recommended for the evaluation of printed or audiovisual materials, we decided to use it as it is a systematic method designed to determine whether patients will be able to understand and act on the information provided. According to the developers' website, materials are understandable when consumers can process and explain key messages regardless of their backgrounds and level of health literacy. On the other hand, materials are actionable when patients can identify what they can do based solely on the information provided, regardless of their background and health literacy level. There are two versions of the PEMAT, for printable and for audiovisual materials. We used the printable version, which consists of 17 understandability items and 7 actionability items. Each item was rated as agree (1 point), disagree (0 points), or N/A (not applicable). The sum of the total points was divided by the total possible points (excluding items with N/A), and the result was multiplied by 100. A higher score indicates that the material is more understandable or actionable. We gave an understandability and actionability score to every response given by the LLMs.

2.4. Statistical Analysis

We calculated and charted the mean, mode, standard deviation (SD), and range of the evaluated metrics of the models' responses using a Microsoft Excel spreadsheet (Version 2403 Build 16.0.17425.20236) 64-bit). To compare the models' performance, we employed the analyses of variance (ANOVA) and Tukey's post hoc analysis when applicable. ANOVA and Tukey's post hoc test were calculated using Microsoft Excel's statistical package. We considered a p -value < 0.05 as statistically significant.

3. Results

3.1. Medical Accuracy

Overall, the three LLMs provided accurate information, with no statistically significant difference ($p = 0.85$). ChatGPT-3.5 obtained the highest mean score of 4.19 ± 0.93 , followed by GPT-4 with a mean of 4.16 ± 0.88 and Gemini with a mean of 4.06 ± 0.91 . The three models' scores ranged between 2 and 5 points, with 81% of the answers provided by ChatGPT-4 scoring higher than 4. The same was true for 78% and 75% of ChatGPT-3.5's and Gemini's answers, respectively (Figure 2).

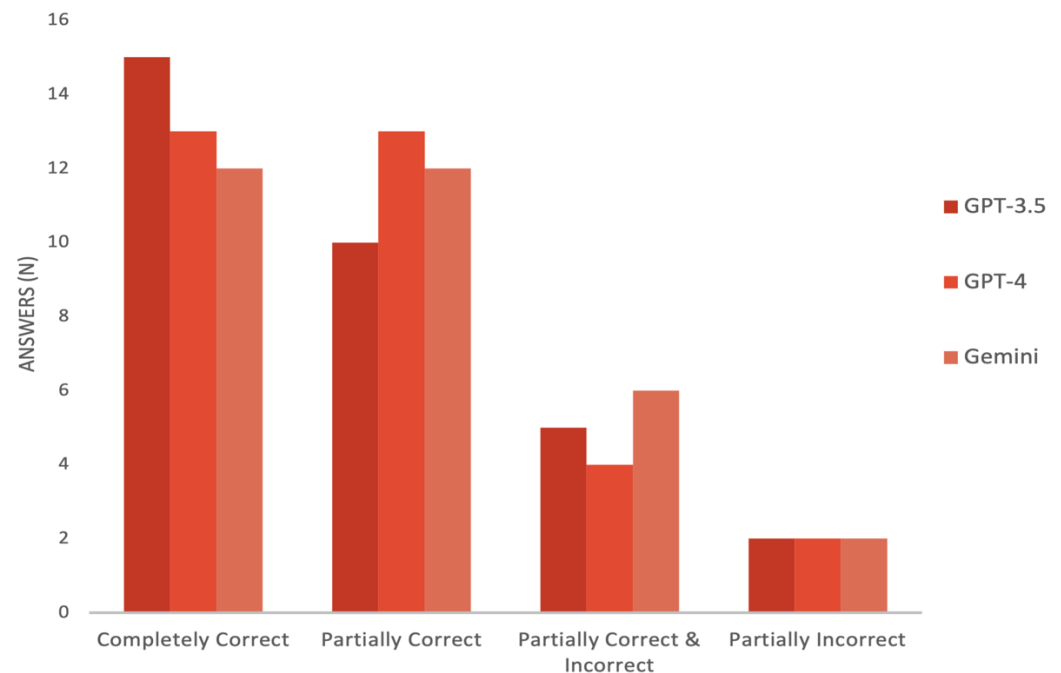


Figure 2. Likert scale scores per LLM.

3.2. Readability

Gemini's responses were significantly more readable than those of ChatGPT-3.5 and 4. The average FRE score for Gemini was 43.72 ± 10.2 , which was significantly higher than ChatGPT-3.5's 33.7 ± 6.8 and ChatGPT-4's 33.7 ± 6.2 ($p = 0.001$). This translated to a significantly lower FKGL average for Gemini's responses (10.92 ± 2.0) than those of ChatGPT-3.5 (12.88 ± 1.0) and ChatGPT-4 (13.6 ± 1.3), with a p -value of 0.001. While 53% of the responses provided by Gemini required a college reading level, it was the only model with responses requiring a 10th to 12th-grade reading level (31%). In contrast, almost 72% of ChatGPT-4 responses required a college reading level, and 41% of those from ChatGPT-3.5 were nearly at a college graduate reading level (Figure 3).

3.3. Understandability and Actionability

Gemini provided more understandable responses, with an average PEMAT understandability score of $90.97 \pm 3.0\%$. This was statistically different from ChatGPT-4, with an average score of $85.13 \pm 4.9\%$ ($p = 0.001$), and ChatGPT-3.5, which averaged at $88.31 \pm 2.7\%$ ($p = 0.014$). ChatGPT-3.5 also obtained a significantly higher average understandability score than GPT-4, with a p -value of 0.002. Regarding actionability, ChatGPT-4 provided the most actionable responses, with an average of $58.7 \pm 8.7\%$. Both ChatGPT-3.5 and Gemini averaged 57.5%; however, Gemini's responses ranged among higher scores, 40–80%, as compared to ChatGPT-3.5, which scored as low as 20% and only reached a maximum score of 60%. Nevertheless, there was no statistically significant difference in actionability among these models ($p = 0.83$) (Figure 4).

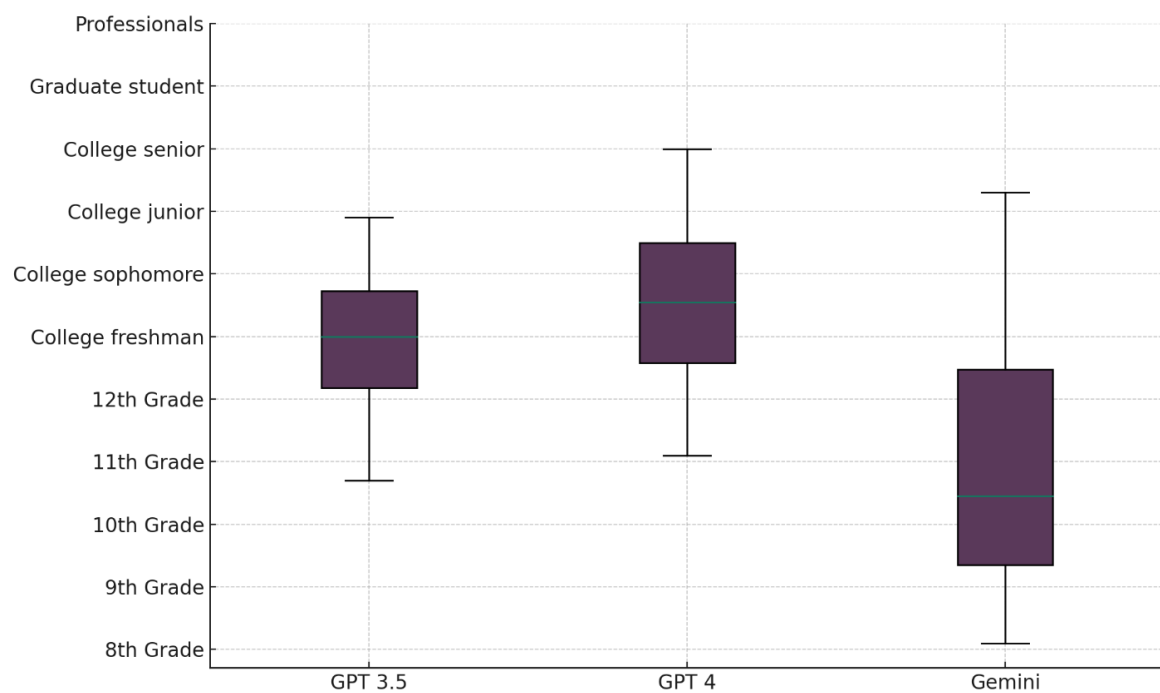


Figure 3. A box and whisker plot of the readability scores per LLM. Each box represents the interquartile range (25–75% of data), the line inside the box shows the median, and the whiskers extend to the smallest and largest values within 1.5 times the interquartile range from the lower and upper quartiles, respectively.

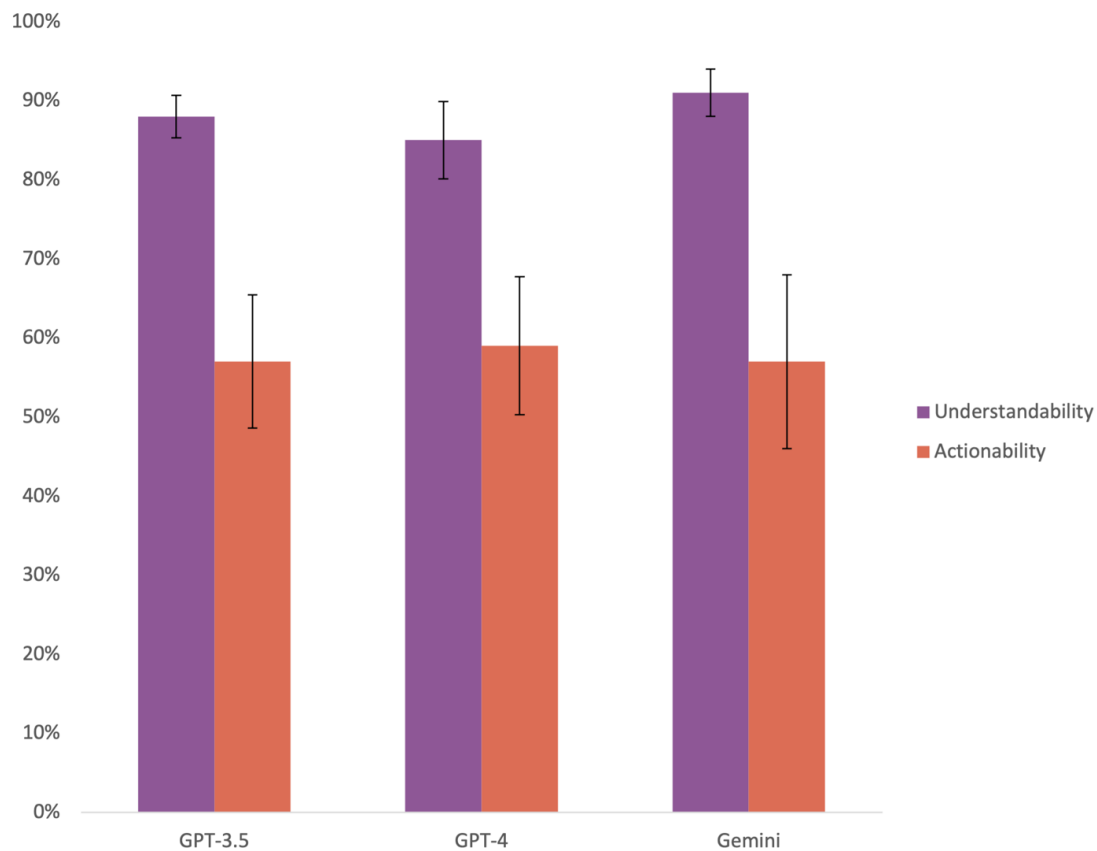


Figure 4. Understandability and actionability scores per LLM. Bars represent average score; error bars represent standard deviation.

4. Discussion

In this modern era, the vast, instantaneous access to medical information acts as a double-edged sword. While patients can easily access helpful information to guide their decisions, they can also easily be misinformed, risking making prejudicial decisions. As they are publicly available, LLMs may present as a solution by offering accurate information presented as a human-like conversation text. One study even showed that patients preferred chatbot responses over physicians' as they were perceived as more empathetic [10]. In specialties with rising demands, such as plastic surgery, these models may be pivotal as they can serve a vast number of patients simultaneously.

In plastic surgery, several studies have analyzed LLMs' capabilities for answering patient questions in the pre- and postoperative period for breast surgery [2,3,8], blepharoplasty [14], rhinoplasty [15], and oculoplastic surgery [28]. This underscores their potential as extremely helpful and valuable adjunctive tools for patient management and the importance of exploring their capabilities, with further visualization toward independent tools (Figure 5). To our knowledge, this is the first study comparing the two versions of ChatGPT and Gemini.



Figure 5. Examples of prompts given to the LLMs and an illustration of the LLMs as adjunctive tools for postoperative care. Created with BioRender [29].

Optimally, LLMs would be able to provide accurate medical information and, when not, at least provide non-harmful advice without over-alerting patients, which would further burden physicians. In our study, the three LLMs provided accurate information at least 75% of the time. Although ChatGPT-3.5 had entirely correct answers almost 47% of the time, it also responded partially incorrectly nearly 22% of the time. The latter was also true for 25% of Gemini's responses. Conversely, ChatGPT-4 was at least partially correct 81% of the time. However, there was no significant difference in the accuracy among the models. This was the opposite in Al-Sharif et al. [28] and in Abi-Rafeh et al. [2], where GPT-3.5 outperformed Bard in providing comprehensive, accurate responses.

Even though there was no statistical difference, we identified that ChatGPT-4's responses were more comprehensive and straightforward than those of ChatGPT-3.5 and Gemini, as they were usually preceded or followed by unnecessary, unuseful content. It was common for all of the models to recommend asking or visiting their surgeon, even when the questions were unrelated to life-threatening scenarios. However, Gemini stated four times that it was unable to provide any medical advice as it was just an LLM and instead encouraged patients to visit a doctor. Similarly, ChatGPT-3.5 started its response by saying it was not a doctor before providing an accurate response twice. This may be an attempt to avoid accountability or a way to express their limitations as LLMs, but it was not the case for ChatGPT-4.

Despite their impressive ability to process information, LLMs still struggle to provide completely accurate responses, which remains the most prevailing concern in their use among specialties [30]. ChatGPT-3.5's latest update was in March 2022, which hinders its ability to answer questions with updated information after that date. Although ChatGPT-4 was last updated in April 2023, the same can still be true. Both GPT-4 and Gemini can access the Internet to provide updated information; nevertheless, their responses are primarily based on their training data [4]. This risks the models being biased as they cannot only inherit but also amplify biases present in their training data [31,32], perpetuating inequalities related to factors such as race, gender, and socioeconomic status [1,11]. Moreover, LLMs may generate fabricated responses, often referred to as hallucinations, when lacking information, which can result in deviations from established practices [30]. Statistical parity ensures that the demographics used for training the models are the same as the demographics of the population where they will be implemented, which can be achieved with specialized training [31,32]. However, for now, it is crucial to analyze inaccurate responses and verify the information generated by the models [30–32].

Our study determined none of the responses as completely incorrect, only partially incorrect. This may be seen in the scenario of an acute postoperative period of a patient who underwent an abdominoplasty and wanted to know how many days she had to stay in bed. ChatGPT-4 suggested gentle walks around the house and no strenuous activities but failed to mention the number of days. On the other hand, Gemini accurately mentioned that it was not recommended to stay in bed as early ambulation was crucial for recovery. However, it did not say why and then contradicted itself by recommending 2 days of strict bed rest followed by short walks and light activity. While low Likert scores do not necessarily mean that the responses might be harmful to the patient, encourage risky behaviors, or provide misinformation, these inaccuracies make them unfit as independent tools.

Although medical accuracy is paramount, LLMs' responses must be readable to successfully serve their purpose as patient resources. The average adult in the United States reads at approximately a 7th-grade reading level [33]. Additionally, one study identified that in plastic surgery, 50.2% of patients had an education level lower than high school and that 48% of attendings, residents, and PAs said it was challenging to make patients follow postoperative instructions [34]. Moreover, patient materials in plastic surgery are often above the recommended average reading level of 6th to 8th grade and may be too difficult for the average patient [34–37].

Vallurupalli et al. [37] successfully used ChatGPT-3.5 to significantly simplify traditional patient education materials for craniofacial surgery by 3 FKGL points. Interestingly,

in our study, ChatGPT-3.5 and 4's average FKGL scores were far superior to the recommended reading level. This was consistent with the results of Momenaei et al. [38], where ChatGPT-4's FKGL and FRE scores averaged at 14.3 and 31.6, respectively. Conversely, Gemini proved to be significantly superior in terms of readability, with responses at an average reading level of 10th grade and an average FRE score 10 points higher than that of the other models. Al-Sharif et al. [28] identified similar results, with GPT having a higher analytical reading inventory (ARI) score than BARD, indicating that a higher level of education was required to understand its responses. Perhaps providing tailored prompts requesting a specific reading level might consistently improve the readability of the answers. However, this might risk oversimplifying the responses and missing essential information.

Gemini also outperformed both ChatGPT models in providing understandable answers. This was because the PEMAT contemplates the use of visual aids as part of its score. Gemini created images for three responses, two of which helped improve understandability. Additionally, in a question about breast augmentation complications, it provided a link to the FDA information website. While ChatGPT-4 also has the ability to create images due to its integration with DALL-E, it only does so when specifically asked to. The ability to think when it would be appropriate to include images for an explanation instead of waiting to be asked to is what determined Gemini's superiority. Nevertheless, both GPT models scored above 85%, proving that they provide understandable answers. Similar to readability, tailoring the prompt so that it provides helpful images may improve further ChatGPT-4's understandability.

Overall, the three models performed poorly in actionability, scoring less than 60% on average. This was a consequence of the format in which the models presented their responses. A minimum of two out of the seven questions were not applicable, and while most of the time they provided clear, broken down actions, they rarely provided tangible tools such as checklists. As visual aids were also part of the actionability section, both GPT models could only score as high as 80%. Although Gemini did provide images, they were not useful for inspiring actionability, hence not showing any difference between the models.

5. Strengths and Limitations

To our knowledge, this is the first study comparing the two versions of ChatGPT and Gemini in terms of providing postoperative care recommendations to patients who have undergone cosmetic plastic surgery procedures. However, our study has some limitations. First are the limited number of questions prompted per procedure. This limited the depth in which we could analyze the models' capacity to provide accurate, readable, understandable, and actionable responses for any particular procedure, limiting their clinical applicability. Furthermore, our findings may not generalize to other LLMs or fields within or outside of plastic surgery and are limited by the current training data of the models. Further research into the newest, more potent versions of Gemini may be paramount, especially considering the current version's superiority to ChatGPT models. Moreover, with additional research constraining the LLMs to provide good-quality information, either through RAG, functional tuning, prompt engineering, or parameterization, we can leverage language understanding but restrict the answers to good, accurate information. Notably, as LLMs evolve rapidly and constantly, our results will likely differ from those obtained in the near future. Nevertheless, a continuous evaluation of model performance is crucial at all stages of development. The limitations and weaknesses highlighted by our study can provide valuable insights to guide future development and practice, especially toward specialty-specific models.

The present study evaluated the LLMs' responses based on their textual contents without considering the patient–physician interaction and feedback. The impact on patient understanding, satisfaction, subsequent behavior, and real-life clinical settings remains unexplored. Lastly, our study primarily focused on English-language responses, demanding further research on the performance across different languages and cultural contexts.

6. Conclusions

Our study provides valuable insights into the efficacy of LLMs in delivering postoperative care recommendations to patients who have undergone cosmetic plastic surgery. Although ChatGPT-4 is the most effective and latest updated version of OpenAI, the publicly available 3.5 version and Google's Gemini public version provided equally accurate medical advice. However, Gemini proved to be superior in providing understandable and more readable responses. While all three models demonstrated their potential as adjunctive tools in postoperative patient care, their shortcomings in providing actionable and concise guidance highlight the need for additional refinement and research to enable their evolution into comprehensive standalone resources. Further experimentation is necessary to evaluate the performance of a retrieval-augmented generation technique.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/healthcare12111083/s1>, Supplementary File S1. Questions provided to LLMs; Supplementary File S2. LLMs Responses.

Author Contributions: Conceptualization, A.J.F., A.S. and C.A.G.-C.; methodology, C.A.G.-C. and S.B.; software, C.A.G.-C. and S.A.H.; formal analysis, C.A.G.-C., S.M.P. and S.B.; validation, S.A.H., S.B. and S.M.P.; investigation, C.A.G.-C., S.M.P. and S.A.H.; writing—original draft, C.A.G.-C. and S.B.; writing—review and editing, S.B., S.M.P., S.A.H., B.C.L., A.S. and A.J.F.; supervision, A.J.F., A.S. and B.C.L.; project administration, A.J.F., A.S. and B.C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Hadi, M.U.; Al-Tashi, Q.; Qureshi, R.; Shah, A.; Muneer, A.; Irfan, M.; Zafar, A.; Shaikh, M.B.; Akhtar, N.; Al-Garadi, M.A.; et al. Large Language Models: A Comprehensive Survey of Applications, Challenges, Limitations, and Future Prospects. *TechRxiv* **2023**. [CrossRef]
2. Abi-Rafeh, J.; Henry, N.; Xu, H.H.; Bassiri-Tehrani, B.; Arezki, A.; Kazan, R.; Gilardino, M.S.; Nahai, F. Utility and Comparative Performance of Current Artificial Intelligence Large Language Models as Postoperative Medical Support Chatbots in Aesthetic Surgery. *Aesthet. Surg. J.* **2024**, sjae025. [CrossRef] [PubMed]
3. Seth, I.; Cox, A.; Xie, Y.; Bulloch, G.; Hunter-Smith, D.J.; Rozen, W.M.; Ross, R.J. Evaluating Chatbot Efficacy for Answering Frequently Asked Questions in Plastic Surgery: A ChatGPT Case Study Focused on Breast Augmentation. *Aesthet. Surg. J.* **2023**, 43, 1126–1135. [CrossRef] [PubMed]
4. Gupta, R.; Pande, P.; Herzog, I.; Weisberger, J.; Chao, J.; Chaiyasate, K.; Lee, E.S. Application of ChatGPT in Cosmetic Plastic Surgery: Ally or Antagonist? *Aesthet. Surg. J.* **2023**, 43, NP587–NP590. [CrossRef] [PubMed]
5. American Society of Plastic Surgeons. American Society of Plastic Surgeons Reveals 2022's Most Sought-After Procedures. Available online: <https://www.plasticsurgery.org/news/press-releases/american-society-of-plastic-surgeons-reveals-2022s-most-sought-after-procedures> (accessed on 12 February 2024).
6. Humar, P.; Asaad, M.; Bengur, F.B.; Nguyen, V. ChatGPT Is Equivalent to First-Year Plastic Surgery Residents: Evaluation of ChatGPT on the Plastic Surgery In-Service Examination. *Aesthet. Surg. J.* **2023**, 43, NP1085–NP1089. [CrossRef] [PubMed]
7. Pan, W.; Liu, D.; Fang, J. An Examination of Factors Contributing to the Acceptance of Online Health Misinformation. *Front. Psychol.* **2021**, 12, 630268. [CrossRef] [PubMed]
8. Yun, J.Y.; Kim, D.J.; Lee, N.; Kim, E.K. A comprehensive evaluation of ChatGPT consultation quality for augmentation mam-moplasty: A comparative analysis between plastic surgeons and laypersons. *Int. J. Med. Inform.* **2023**, 179, 105219. [CrossRef] [PubMed]
9. Aliyeva, A.; Sari, E.; Alaskarov, E.; Nasirov, R. Enhancing Postoperative Cochlear Implant Care With ChatGPT-4: A Study on Artificial Intelligence (AI)-Assisted Patient Education and Support. *Cureus* **2024**, 16, e53897. [CrossRef] [PubMed]

10. Ayers, J.W.; Poliak, A.; Dredze, M.; Leas, E.C.; Zhu, Z.; Kelley, J.B.; Faix, D.J.; Goodman, A.M.; Longhurst, C.A.; Hogarth, M.; et al. Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA Intern. Med.* **2023**, *183*, 589–596. [CrossRef]
11. Clusmann, J.; Kolbinger, F.R.; Muti, H.S.; Carrero, Z.I.; Eckardt, J.N.; Laleh, N.G.; Löffler, C.M.L.; Schwarzkopf, S.C.; Unger, M.; Veldhuizen, G.P.; et al. The future landscape of large language models in medicine. *Commun. Med.* **2023**, *3*, 141. [CrossRef]
12. Dias, P.; Clerc, D.; da Rocha Rodrigues, M.G.; Demartines, N.; Grass, F.; Hubner, M. Impact of an Operating Room Nurse Preoperative Dialogue on Anxiety, Satisfaction and Early Postoperative Outcomes in Patients Undergoing Major Visceral Surgery—A Single Center, Open-Label, Randomized Controlled Trial. *J. Clin. Med.* **2022**, *11*, 1895. [CrossRef] [PubMed]
13. Putterman, A.M. Patient satisfaction in oculoplastic surgery. *Ophthalmic Surg. Lasers Imaging Retin.* **1990**, *21*, 15–21. [CrossRef]
14. Cox, A.; Seth, I.; Xie, Y.; Hunter-Smith, D.J.; Rozen, W.M. Utilizing ChatGPT-4 for Providing Medical Information on Blepharoplasties to Patients. *Aesthet. Surg. J.* **2023**, *43*, NP658–NP662. [CrossRef] [PubMed]
15. Soto-Galindo, G.A.; Capelleras, M.; Cruellas, M.; Apaydin, F. Effectiveness of ChatGPT in Identifying and Accurately Guiding Patients in Rhinoplasty Complications. *Facial Plast. Surg.* **2023**, online ahead of print. [CrossRef]
16. Atkinson, C.J.; Seth, I.; Xie, Y.; Ross, R.J.; Hunter-Smith, D.J.; Rozen, W.M.; Cuomo, R. Artificial Intelligence Language Model Performance for Rapid Intraoperative Queries in Plastic Surgery: ChatGPT and the Deep Inferior Epigastric Perforator Flap. *J. Clin. Med.* **2024**, *13*, 900. [CrossRef] [PubMed]
17. American Society of Plastic Surgeons. What are the Risks of Eyelid Surgery? Available online: https://www.plasticsurgery.org/cosmetic-procedures/eyelid-surgery/safety?_gl=1*4kj7fz*_ga*OTU4NTgxOTM3LjE3MDc5NDcwOTY.*_ga_2N9PTXXTDN*MTcxMDM0ODc1MC44LjEuMTcxMDM0ODkwOC41Ny4wLjA (accessed on 2 February 2024).
18. American Society of Plastic Surgeons. What are the Risks of Breast Augmentation? Available online: https://www.plasticsurgery.org/cosmetic-procedures/breast-augmentation/safety?_gl=1*aa45id*_ga*OTU4NTgxOTM3LjE3MDc5NDcwOTY.*_ga_2N9PTXXTDN*MTcxMDk0NjQwOC4xMS4xLjE3MTA5NDY2OTEuMTYuMC4w%20 (accessed on 2 February 2024).
19. American Society of Plastic Surgeons. What Are the Risks of Breast Lift Surgery? Available online: https://www.plasticsurgery.org/cosmetic-procedures/breast-lift/safety?_gl=1*1xixp8*_ga*OTU4NTgxOTM3LjE3MDc5NDcwOTY.*_ga_2N9PTXXTDN*MTcxMDk0NjQwOC4xMS4xLjE3MTA5NDY4MDEuNTkuMC4w (accessed on 2 February 2024).
20. American Society of Plastic Surgeons. What Are the Risks of Liposuction? Available online: https://www.plasticsurgery.org/cosmetic-procedures/liposuction/safety?_gl=1*xsqaqz*_ga*OTU4NTgxOTM3LjE3MDc5NDcwOTY.*_ga_2N9PTXXTDN*MTcxMDk0NjQwOC4xMS4xLjE3MTA5NDY5MjcuNTkuMC4w (accessed on 2 February 2024).
21. American Society of Plastic Surgeons. What Are the Risks of Tummy Tuck Surgery? Available online: https://www.plasticsurgery.org/cosmetic-procedures/tummy-tuck/safety?_gl=1*n5pmbf*_ga*OTU4NTgxOTM3LjE3MDc5NDcwOTY.*_ga_2N9PTXXTDN*MTcxMDk0NjQwOC4xMS4xLjE3MTA5NDcwODYuNTIuMC4w (accessed on 2 February 2024).
22. *The Art of Aesthetic Surgery: Principles & Techniques*, 3rd ed.; Nahai, F.; Nahai, F. (Eds.) Thieme: New York, NY, USA, 2020. [CrossRef]
23. *Essentials of Plastic Surgery*, 3rd ed.; Janis, J.E. (Ed.) Thieme: New York, NY, USA, 2022.
24. Readable. Flesch Reading Ease and the Flesch Kincaid Grade Level. Available online: <https://readable.com/readability/flesch-reading-ease-flesch-kincaid-grade-level/> (accessed on 12 February 2024).
25. Weis, B. *Health Literacy: A Manual for Clinicians*; American Medical Association, American Medical Foundation: Chicago, IL, USA, 2003.
26. Rooney, M.K.; Santiago, G.; Perni, S.; Horowitz, D.P.; McCall, A.R.; Einstein, A.J.; Jagsi, R.; Golden, D.W. Readability of Patient Education Materials From High-Impact Medical Journals: A 20-Year Analysis. *J. Patient Exp.* **2021**, *8*, 2374373521998847. [CrossRef] [PubMed]
27. Shoemaker, S.J.; Wolf, M.S.; Brach, C. Development of the Patient Education Materials Assessment Tool (PEMAT): A new measure of understandability and actionability for print and audiovisual patient information. *Patient Educ. Couns.* **2014**, *96*, 395–403. [CrossRef] [PubMed]
28. Al-Sharif, E.M.; Penteado, R.C.; Dib El Jalbout, N.; Topilow, N.J.; Shoji, M.K.; Kikkawa, D.O.; Liu, C.Y.; Korn, B.S. Evaluating the Accuracy of ChatGPT and Google BARD in Fielding Oculoplastic Patient Queries: A Comparative Study on Artificial versus Human Intelligence. *Ophthalmic Plast. Reconstr. Surg.* **2024**, *40*, 303–311. [CrossRef] [PubMed]
29. BioRender. Available online: <https://www.biorender.com/> (accessed on 16 February 2024).
30. Pressman, S.M.; Borna, S.; Gomez-Cabello, C.A.; Haider, S.A.; Haider, C.; Forte, A.J. AI and Ethics: A Systematic Review of the Ethical Considerations of Large Language Model Use in Surgery Research. *Healthcare* **2024**, *12*, 825. [CrossRef]
31. Leslie, D.; Mazumder, A.; Peppin, A.; Wolters, M.K.; Hagerty, A. Does “AI” stand for augmenting inequality in the era of covid-19 healthcare? *BMJ* **2021**, *372*, n304. [CrossRef]
32. Zaidi, D.; Miller, T. Implicit Bias and Machine Learning in Health Care. *South. Med. J.* **2023**, *116*, 62–64. [CrossRef] [PubMed]
33. Ricci, J.A.; Vargas, C.R.; Chuang, D.J.; Lin, S.J.; Lee, B.T. Readability assessment of online patient resources for breast augmentation surgery. *Plast. Reconstr. Surg.* **2015**, *135*, 1573–1579. [CrossRef] [PubMed]
34. Barton, N.; Janis, J.E. Missing the Mark: The State of Health Care Literacy in Plastic Surgery. *Plast. Reconstr. Surg. Glob. Open* **2020**, *8*, e2856. [CrossRef] [PubMed]
35. Chen, A.D.; Ruan, Q.Z.; Bucknor, A.; Chattha, A.S.; Bletsis, P.P.; Furnas, H.J.; Lee, B.T.; Lin, S.J. Social Media: Is the Message Reaching the Plastic Surgery Audience? *Plast. Reconstr. Surg.* **2019**, *144*, 773–781. [CrossRef] [PubMed]

36. Fanning, J.E.; Okamoto, L.A.; Levine, E.C.; McGee, S.A.; Janis, J.E. Content and Readability of Online Recommendations for Breast Implant Size Selection. *Plast. Reconstr. Surg. Glob. Open* **2023**, *11*, e4787. [[CrossRef](#)] [[PubMed](#)]
37. Vallurupalli, M.; Shah, N.D.; Vyas, R.M. Validation of ChatGPT 3.5 as a Tool to Optimize Readability of Patient-facing Craniofacial Education Materials. *Plast. Reconstr. Surg. Glob. Open* **2024**, *12*, e5575. [[CrossRef](#)]
38. Momenaei, B.; Wakabayashi, T.; Shahlaee, A.; Durrani, A.F.; Pandit, S.A.; Wang, K.; Mansour, H.A.; Abishek, R.M.; Xu, D.; Sridhar, J.; et al. Appropriateness and Readability of ChatGPT-4-Generated Responses for Surgical Treatment of Retinal Diseases. *Ophthalmol. Retin.* **2023**, *7*, 862–868. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.